

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

Maucha Andrade Gamonal

**COPA 2014 FRAMENET BRASIL: DIRETRIZES PARA A CONSTITUIÇÃO DE UM
DICIONÁRIO ELETRÔNICO TRILÍNGUE A PARTIR DA ANÁLISE DE FRAMES DA
EXPERIÊNCIA TURÍSTICA**

Juiz de Fora
2013

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

**COPA 2014 FRAMENET BRASIL: DIRETRIZES PARA A CONSTITUIÇÃO DE UM
DICIONÁRIO ELETRÔNICO TRILÍNGUE A PARTIR DA ANÁLISE DE FRAMES DA
EXPERIÊNCIA TURÍSTICA**

Maucha Andrade Gamonal

Dissertação de Mestrado apresentada ao programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como parte dos requisitos necessários à obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Tiago Timponi Torrent

Juiz de Fora
Março de 2013

**COPA 2014 FRAMENET BRASIL: DIRETRIZES PARA A CONSTITUIÇÃO DE UM
DICIONÁRIO ELETRÔNICO TRILÍNGUE A PARTIR DA ANÁLISE DE FRAMES DA
EXPERIÊNCIA TURÍSTICA**

Maucha Andrade Gamonal

Orientador: Prof. Dr. Tiago Timponi Torrent

Dissertação de Mestrado submetida ao programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como parte dos requisitos necessários à obtenção do título de Mestre em Linguística.

Aprovada por:

Presidente, Prof. Dr. Tiago Timponi Torrent – UFJF

Prof. Dr. Bento Carlos Dias da Silva – UNESP/Araraquara

Profa. Dra. Maria Margarida Martins Salomão – UFJF

Juiz de Fora
Março de 2013

O significado não é algo em si; ele envolve o que é significativo para nós. Nada é significativo em si mesmo. O significado deriva da experiência de atuação como um ser de certo tipo em um ambiente de certo tipo.

George Lakoff

A minha Vó Lizete, por me ensinar coisas que eu jamais aprenderia com os livros.

AGRADECIMENTOS

O desenvolvimento da ciência é um trabalho solitário e, ao mesmo tempo, coletivo. Solitário porque boa parte da contribuição a ser apresentada requer reflexão, leituras, pesquisas, análises, textos, reescritas... É coletivo porque conhecimento não se constrói individualmente, as tarefas acima citadas estão em constante diálogo com outras propostas, outras teorias, outros olhares e outros posicionamentos. Dessa forma, a tentativa de agradecer pessoas e entidades que fizeram parte deste trabalho segue sendo apenas uma tentativa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Capes, por ter garantido, com a concessão da bolsa de mestrado, as condições materiais para a realização deste trabalho.

A toda equipe do Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora, a atenção, a disposição e a base acadêmica foram fundamentais para a minha formação.

Ao Ministério do Turismo brasileiro e aos autores dos blogs de viajantes. A disponibilização dos textos foi imprescindível para as análises realizadas neste trabalho.

Ao meu orientador professor doutor Tiago Timponi Torrent, obrigada pelas orientações atenciosas e pela paciência em todos os momentos.

Aos professores que se dispuseram a compor a banca avaliadora deste trabalho, professor doutor Bento Carlos Dias da Silva e professora doutora Margarida Salomão. As contribuições dos senhores serão valiosas para a conclusão deste trabalho.

À pesquisadora da FrameNet de Berkeley, Miriam Petruck, os conselhos e as intervenções guiaram a análise de dados desta dissertação.

À professora Palmira Marrafa, da Universidade de Lisboa, por ter, prontamente, disponibilizado leituras e aplicativos da WordNet desenvolvida em Portugal.

Aos pesquisadores da FrameNet Brasil, graduandos, mestrandos, doutorandos e professores. Os encontros e as conversas foram essenciais para este trabalho. Em especial, às graduandas Carolina Alcântara, Élide Costa e Isabela Dutra, que muito colaboraram, neste último período, na coleta de *corpora* e na

análise de dados. Também tenho de agradecer, com carinho especial, ao Ely Matos, nosso parceiro da computação. Nossas conversas foram fontes de conhecimento.

À professora doutora Ana Claudia Peters. Suas considerações em classe, ponderando a diferença entre nós, alunos, e vocês, professores, foram essenciais para a maturidade desta proposta.

Aos colegas da Faculdade de Letras da UFJF. O papel de representante discente, durante os anos da graduação, permitiu-me acumular muitas experiências, e, hoje, eu as reconheço nas linhas desta dissertação.

Ao amigo Fábio Duque, a fonte de bibliografia fornecida foi dando forma a este trabalho e me possibilitou economizar muitos dólares.

À família. Minha mãe, Vanda Lea, minha Vó Lizete, tias Wânia Andrade e Milca Andrade e meu irmão, Fagner Gamonal, agradeço a vocês em nome de toda a família. O suporte, carinho e paciência fazem de vocês o pilar da minha vida. Estendo esse agradecimento à família Sales e à Dona Ilda Nader pelo carinho e torcida sempre.

Aos amigos. Sem vocês, a vida não seria a mesma coisa. Agradeço a vocês (inclusive aos ciumentos Priscilla Chandretti e Luã Cupolillo) em nome de um grande amigo, Vitor Gonçalves. Sua amizade é meu protótipo de amigo. Aquele que não precisa te ver todos os dias para saber que está lá, disponível para o que der e vier. Obrigada!

Ao Hyllo, meu grande amor! Obrigada por ser tão companheiro. Dividir a vida com você é fazer valer a pena todos os segundos.

À Vida, pela vontade de aprender, de ensinar e de aprender ensinando.

Gamonal, Maucha Andrade.

COPA 2014 FRAMENET BRASIL: DIRETRIZES PARA A CONSTITUIÇÃO DE UM DICIONÁRIO ELETRÔNICO TRILÍNGUE A PARTIR DA ANÁLISE DE FRAMES DA EXPERIÊNCIA TURÍSTICA / Maucha Andrade Gamonal. -- 2013.

145 f. : il.

Orientador: Tiago Timponi Torrent

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2013.

1. Semântica de Frames. 2. FrameNet Brasil. 3. Dicionários Eletrônicos Multilíngues. 4. Turismo. 5. Copa do Mundo FIFA 2014. I. Torrent, Tiago Timponi, orient. II. Título.

RESUMO

Esta dissertação é parte do subprojeto Copa 2014 FrameNet Brasil (SALOMÃO ET AL., 2011), iniciativa da FrameNet Brasil em parceria com o projeto FrameCorp (CHISHMAN ET AL., 2008) e com a Berkeley FrameNet (FILLMORE ET AL., 2003) que propõe a elaboração de dicionário eletrônico trilingue – Português, Inglês, Espanhol – para os domínios da Copa do Mundo, do Futebol e do Turismo. A elaboração deste recurso se mostra diferente dos demais dicionários eletrônicos por ser estruturado a partir da teoria da Semântica de *Frames* (FILLMORE, 1982, 1985; PETRUCK, 1996) e da metodologia da FrameNet (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010). A contribuição desta pesquisa é estabelecer as diretrizes para a estruturação deste dicionário a partir da modelagem de *frames* da experiência turística. Para tanto, algumas perguntas guiam o desenvolvimento deste trabalho: i) Em que medida os *frames* do domínio turístico modelados com *corpora* compilados da língua portuguesa do Brasil servem para representar os *frames* do Turismo para as demais línguas do dicionário? ii) Como a FrameNet responde aos desafios colocados na estruturação de recursos lexicais multilíngues? É possível utilizar *frames* como interlíngua? iii) Que avaliação se pode fazer do Kicktionary, dicionário multilíngue do futebol, como produto que utiliza a FrameNet e a WordNet (MILLER, 1993, 1995; FELLBAUM, 1998)? As respostas a essas reflexões apontaram que: i) os *frames* do domínio turístico são modelados da mesma forma pelas diferentes culturas; ii) a rede semântica FrameNet precisa se adaptar às especificidades impostas pela lexicografia multilíngue, mas, devido ao caráter transcultural do Turismo e também da Copa do Mundo e do Futebol, os *frames* podem ser utilizados como interlíngua; iii) o Kicktionary, como dicionário multilíngue de domínio especializado que utiliza *frames* e *synsets*, deveria rever a funcionalidade de cada teoria na estruturação dos bancos de dados. As relações intralinguísticas poderiam acontecer via WordNet, e as relações interlinguísticas, via FrameNet.

Palavras-chave: Semântica de Frames; FrameNet Brasil; Dicionários Eletrônicos Multilíngues; Turismo; Copa do Mundo FIFA 2014.

ABSTRACT

This work is part of the subproject 2014 World Cup FrameNet Brasil (SALOMÃO ET AL., 2011), an initiative of FrameNet Brazil in partnership with the FrameCorp project (CHISHMAN ET AL., 2008) and Berkeley FrameNet (FILLMORE ET AL. 2003), which proposes developing a trilingual electronic dictionary - English, Portuguese, Spanish - for the domains of the World Cup, Soccer, and Tourism. The development of this resource is different from the other electronic dictionaries because it is structured within the theory of Frame Semantics (FILLMORE, 1982, 1985; PETRUCK, 1996) and the methodology of FrameNet (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010). The contribution of this research is to establish guidelines for structuring this dictionary through the analysis of tourist experience frames. Therefore, some questions guide the development of this work: i) How does this resource respond to the challenges involved in structuring multilingual lexical resources? ii) Is it possible to use frames as an interlingual representation? iii) What evaluation can be made of Kicktionary, a multilingual dictionary of football as a product that uses both FrameNet and WordNet (Miller, 1993, 1995; FELLBAUM, 1998)? The answers to these considerations have shown that: i) frames of the tourist domain are modeled in the same way by different cultures; ii) FrameNet needs to adapt to the specificities imposed by multilingual lexicography, but due to the transcultural nature of Tourism and also the Soccer and the World Cup, frames can be used as interlingua; iii) the Kicktionary, as specialized multilingual dictionary which uses frames and synsets, should review the functionality of each theory in structuring databases. Intralinguistic relations could happen via WordNet, and interlingual relations via FrameNet.

Keywords: Frame Semantics; FrameNet Brasil; Multilingual Electronic Dictionaries; Tourism; FIFA 2014 World Cup.

LISTA DE FIGURAS

Figura 1: Texto em formato <i>.txt</i>	25
Figura 2: Etapas do tratamento computacional	26
Figura 3: Resultado da busca pelo lema <i>turista</i> na ferramenta <i>Concordance</i>	27
Figura 4: Resultado da busca por <i>turista</i> na ferramenta <i>Word Sketch</i>	28
Figura 5: <i>Corpora</i> no <i>software</i> FrameNet Desktop	29
Figura 6: Processo de anotação em três camadas na FrameNet Brasil	32
Figura 7: Dicionário eletrônico <i>offline</i> e <i>online</i>	35
Figura 8: Busca por <i>apreciar</i> na seção do dicionário analógico no Dicionário Criativo	36
Figura 9: Resultado da busca por <i>travel</i> na ferramenta <i>FrameNet Search</i>	47
Figura 10: Resultado da busca por <i>Travel</i> na ferramenta <i>Frame Index</i>	48
Figura 11: Resultado da busca por <i>travel</i> na ferramenta <i>Lexical Unit Index</i>	49
Figura 12: Resultado da busca por <i>travel</i> na opção <i>Lexical Entry</i>	49
Figura 13: Resultado da busca por <i>travel</i> na opção <i>Annotation</i>	50
Figura 14: Resultado da busca por <i>Travel</i> na ferramenta <i>FrameGrapher</i>	51
Figura 15: Relação entre os Elementos do <i>Frame</i> <i>Self_motion</i> e <i>Travel</i>	51
Figura 16: Definição do <i>frame</i> <i>Travel</i>	52
Figura 17: Definição dos Elementos de <i>Frame</i> nucleares em <i>Travel</i>	54
Figura 18: Definição dos Elementos de <i>Frame</i> não nucleares em <i>Travel</i>	55
Figura 19: Anotação de Texto Corrido: História de Las Vegas.....	60
Figura 20: Anotação lexicográfica: sentenças que ilustram a UL <i>viajar</i> no <i>frame</i> <i>Travel</i>	62
Figura 21: Anotação da Unidade Lexical <i>viajar</i> na FrameNet Brasil	63
Figura 22: Anotação de conflação de Elemento de <i>Frame</i>	63

Figura 23: Anotação de incorporação de Elemento de <i>Frame</i>	64
Figura 24: Anotação semântica e sintática com a UL <i>viajar</i>	65
Figura 25: Anotação de verbo suporte	65
Figura 26: <i>Frame</i> <code>Employment_scenario</code>	66
Figura 27: Legenda de relações entre <i>frames</i>	66
Figura 28: Unidade Lexical <i>llegar</i> na base de dados da Spanish FrameNet.....	72
Figura 29: Busca pelo lema <i>visit</i> na <i>WordNet</i>	766
Figura 30: Estrutura de adjetivos bipolares	79
Figura 31: <i>Inter-lingual-Index</i> na base de dados da <i>EuroWordNet</i>	82
Figura 32: Lextec - Léxico Técnico do Português	83
Figura 33: Resultado da busca pelo lema <i>bola</i> no <i>TemaNet</i>	83
Figura 34: Visualização de <i>On the Pitch</i> no <i>Kicktionary</i>	85
Figura 35: Unidade Lexical do <i>frame</i> <code>Challenge</code>	86
Figura 36: Caracterização informal dos conceitos do turismo	90
Figura 37: Definição do <i>Frame</i> <code>Touring</code> e seus elementos nucleares.....	92
Figura 38: <code>Visiting_scenario</code>	93
Figura 39: <code>Cenário_do_turismo</code>	96
Figura 40: Sentença com UL <i>conhecer</i>	100
Figura 41: Sentença com UL <i>atrair</i>	100
Figura 42: Sentença com UL <i>abrigar</i>	100
Figura 43: Níveis de profundidade do conhecimento nos sistemas de tradução automática.....	105

LISTA DE QUADROS

Quadro 1: <i>Corpora</i> compilados para o domínio do Turismo.....	23
Quadro 2: Realizações de Elementos de <i>Frame</i> na UL <i>employ</i>	69
Quadro 3: Padrões sintáticos dos Elementos de <i>Frame</i> na UL <i>employ</i>	70
Quadro 4: <i>Frames</i> da Experiência Turística – Chegada	99
Quadro 5: <i>Frames</i> da Experiência Turística – Estada.....	102
Quadro 6: <i>Frames</i> da Experiência Turística – Partida.....	103
Quadro 7: Realização semântica de <i>aproximarse</i> e <i>approach</i>	112

SUMÁRIO

LISTA DE FIGURAS	xi
LISTA DE QUADROS	xiii
INTRODUÇÃO	16
1. METODOLOGIA	19
1.1. Constituição de Corpora para Dicionários Eletrônicos Temáticos	19
1.1.1. <i>Compilação dos Corpora</i>	20
1.1.2. <i>Tratamento Computacional</i>	25
1.2. Métodos para Estruturação de Frames	30
1.3. Anotação Lexicográfica na FrameNet Brasil	31
2. RECURSOS LEXICAIS ELETRÔNICOS PARA USO HUMANO E PARA APRENDIZADO DE MÁQUINAS	33
2.1. Dicionários Eletrônicos	33
2.1.1. <i>Dicionários Impressos vs. Dicionários Eletrônicos</i>	37
2.1.2. <i>Dicionários Eletrônicos Multilíngues</i>	39
2.2. A FrameNet	41
2.2.1. <i>A Semântica de Frames nos estudos cognitivos da linguagem</i>	41
2.2.2. <i>Desenvolvimento da Rede Semântica</i>	46
2.2.3. <i>Expansão da FrameNet para Outras Línguas</i>	71
2.3. A WordNet	74
2.3.1. <i>O Léxico na WordNet</i>	74
2.3.2. <i>Expansão da WordNet para Outras Línguas</i>	81
2.4. O Kicktionary: Um Dicionário Multilíngue para O Futebol	84
3. O CENÁRIO DO TURISMO	88
3.1. O Turismo como Atividade Humana	88
3.2. Criação do Cenário do Turismo para a FrameNet	91
3.2.1. <i>Frames de Cenário</i>	93
3.2.2. <i>Cenário_do_Turismo</i>	94
3.3. Frames da Experiência Turística	98

3.3.1. <i>Frames do Cenário do Turismo Chegada</i>	98
3.3.2. <i>Frames do Cenário do Turismo Estada</i>	99
3.3.3. <i>Frames do Cenário do Turismo Partida</i>	102
4. CONTRIBUIÇÕES DA MODELAGEM DO CENÁRIO DO TURISMO PARA O DESENVOLVIMENTO DO COPA 2014	104
4.1. Frames como Interlíngua	104
4.2. Proposta de Revisão da Estruturação do Kicktionary	113
4.3. O Papel da WordNet no Copa 2014	117
CONCLUSÕES	119
REFERÊNCIAS BIBLIOGRÁFICAS	123
ANEXO	127

INTRODUÇÃO

A FrameNet Brasil – FN Br – é a contraparte para o português do Brasil da rede semântica FrameNet (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010), projeto em desenvolvimento desde o ano de 1997 no International Computer Science Institute, em Berkeley, Califórnia. Idealizada e desenvolvida pelo professor Charles J. Fillmore e sua equipe, a FrameNet propõe a criação de recurso lexical *online*, baseado em informação empírica, descrito nos moldes da teoria da Semântica de *Frames* (FILLMORE, 1982, 1985; PETRUCK, 1996).

Esta dissertação se caracteriza como parte do subprojeto Copa 2014 FrameNet Brasil (SALOMÃO ET AL., 2011), doravante Copa 2014, iniciativa da FN Br em parceria com o projeto FrameCorp (CHISHMAN ET AL., 2008) e com a Berkeley FrameNet, que propõe a elaboração de um dicionário eletrônico trilingue – Português, Inglês, Espanhol com os domínios da Copa do Mundo, do Futebol e do Turismo, organizado a partir de *frames*, estruturas de conhecimento que preveem que a compreensão de um item lexical ou de uma experiência se subordina à compreensão de todo o sistema de conceitos ao qual esse item lexical ou essa experiência esteja vinculado.

Uma demonstração de dicionário multilíngue organizado em *frames* é o Kicktionary (SCHMIDT, 2006, 2007, 2008, 2009). Esse recurso lexical estrutura o vocabulário do futebol com o aporte teórico da Semântica de *Frames* e da WordNet (MILLER, 1993, 1995; FELLBAUM, 1998), rede semântica que organiza o léxico por classes de palavras lexicais através das relações de sentido, organizadas em conjuntos de sinônimos, os *synsets*.

Por considerar que os dicionários tradicionais, muitas vezes, falham ao tentar dissociar o conhecimento linguístico do conhecimento enciclopédico, o intuito maior com o desenvolvimento do dicionário Copa 2014 é verificar em que medida a teoria linguística da Semântica de *Frames* pode contribuir para a criação de recursos lexicais multilíngues. Diante desse propósito, o escopo deste trabalho de mestrado é analisar *frames* da experiência turística a partir da estruturação do *frame* `Cenário_do_turismo`.

Nesse sentido, algumas questões surgem como pontos de partida para o desenvolvimento desta pesquisa:

- i) Em que medida os *frames* do domínio turístico modelados com *corpora* compilados da língua portuguesa do Brasil servem para representar os *frames* do Turismo para as demais línguas do dicionário?
- ii) Como a FrameNet responde aos desafios colocados na estruturação de recursos lexicais multilíngues? É possível utilizar *frames* como interlíngua?
- iii) Que avaliação se pode fazer do Kicktionary como produto que utiliza *frames* e *synsets* para o tratamento lexicográfico multilíngue?

Com o desenvolvimento desta pesquisa, foi possível chegar às seguintes considerações: os *frames* do domínio turístico são perspectivados da mesma forma pelas diferentes culturas, assim como a Copa do Mundo e a modalidade de futebol adotada neste evento. Isso assinalou, conforme o questionamento da segunda pergunta, a possibilidade de os *frames* serem utilizados como interlíngua, o que não isenta a necessidade de a FrameNet se adaptar às especificidades impostas pela lexicografia multilíngue, de modo geral. Quanto ao Kicktionary como dicionário multilíngue de domínio especializado que utiliza *frames* e *synsets*, a funcionalidade de cada teoria no dicionário poderia ser revista. As relações intralinguísticas poderiam se dar via WordNet por oferecerem uma organização pormenorizada das relações lexicais e das categorias lexicais das palavras a partir da estrutura dos *synsets*. Já as relações interlinguísticas, via FrameNet, pois permitiria encontrar, através de mapeamento indireto fornecido pela análise de *corpora*, os melhores correspondentes de tradução. Além disso, a melhor alternativa para estruturar o Futebol é lançar mão de *frames* cenários, conforme realizado para o domínio do Turismo, e não a distinção *frames* e cenas. A terminologia adotada por Schmidt considera os *frames* como estruturas linguísticas e as cenas como estruturas de conhecimento independentes de língua, distinção não concebida pela Linguística Cognitiva, por defender que conhecimento linguístico e conhecimento enciclopédico não podem ser separados.

Ao longo deste trabalho, tais questões são estudadas e delineiam diretrizes para o desenvolvimento do dicionário Copa 2014. Além deste capítulo de introdução e do de conclusão, este trabalho contém outros quatro capítulos.

No capítulo 1, é apresentada a metodologia deste trabalho no que diz respeito à constituição de *corpora* para dicionários eletrônicos temáticos, à estruturação de *frames*, bem como ao procedimento de anotação lexicográfica da FrameNet Brasil.

No capítulo 2, a proposta é inserir os *frames* na confecção de dicionários eletrônicos. Para isso, apresentam-se um panorama geral de recursos lexicais em meio eletrônico assim como os desafios para as versões multilíngues. Neste capítulo, dois recursos lexicais são colocados como alternativas lexicais que servem tanto para usuários humanos quanto para o aprendizado de máquinas: a FrameNet e a WordNet. O Kicktionary é ilustrado como proposta concluída de dicionário multilíngue eletrônico de domínio especializado que utiliza os dois recursos.

No capítulo 3, é apresentado o processo de construção e análise dos *frames* da experiência turística. A complexidade do Turismo apontou a necessidade de estruturar um *frame* que fosse capaz de representar a organização dessa atividade bem como o diálogo com outros domínios, o *frame* `Cenário_do_turismo`.

No capítulo 4, especifica-se como a modelagem do Turismo contribuiu para a estruturação do Copa 2014 como um todo. A proposta de utilizar *frames* como interlíngua foi possibilitada devido ao caráter transcultural observado tanto para o domínio do Turismo como para o domínio da Copa e do Futebol. O dicionário Kicktionary atuou no sentido de mostrar como as teorias linguísticas impulsionadas pela FrameNet e pela WordNet poderiam atuar em conjunto no âmbito lexicográfico. O resultado disso foram propostas de revisão da estrutura do Kicktionary e de discussão do papel ideal da WordNet na estruturação do Copa 2014.

1 METODOLOGIA

O propósito deste capítulo é expor a metodologia empregada na realização deste trabalho. Para tanto, apresentam-se, na seção 1.1, os procedimentos realizados para constituição de dicionários eletrônicos temáticos no que tange à constituição de *corpora* e ao tratamento computacional para o processamento dos dados. Nas seções 1.2 e 1.3, os métodos utilizados para a estruturação de *frames* bem como os passos seguidos no processo de anotação lexicográfica na FrameNet Brasil são exibidos.

1.1 Constituição de *Corpora* para Dicionários Eletrônicos Temáticos

Como o dicionário eletrônico Copa 2014 adota a teoria linguística da Semântica de *Frames* (FILLMORE, 1982, 1985), programa que propõe investigar as continuidades existentes entre linguagem e experiência, a necessidade de subsidiar essa pesquisa com informação empírica é de suma importância. Utilizar as ferramentas da Linguística de *Corpus* não significa, entretanto, abandonar as intuições analíticas esperadas de um especialista da linguagem. Fillmore (1992) procurou levantar essa questão ao esboçar o perfil de um linguista que se baseia na introspecção de falante nativo e o daquele que tem em mãos os fatos observáveis ao se manusearem centenas de textos e milhares de palavras. Para o autor, o estudo em *corpus* oferece a chance de pesquisar coisas que o linguista jamais descobriria de outra forma, mas, ainda assim, o ideal é que intuição e evidência sejam traços presentes em um mesmo analista.

O processo de elaboração de dicionários requer um posicionamento crítico e, ao mesmo tempo, intuitivo do lexicógrafo. Ele assume a função de selecionar o léxico que será contemplado, além de especificar os variados sentidos de um lexema, por exemplo. Dessa forma, unir a intuição de falante nativo ou de conhecedor do idioma à informação empírica garante a representatividade do recurso lexical elaborado.

1.1.1 Compilação dos Corpora

O dicionário Copa 2014 está sendo estruturado como recurso eletrônico multilíngue – Português, Inglês, Espanhol –, que abrange três domínios: a Copa do Mundo, o Futebol e o Turismo. Como o enfoque desta dissertação é apresentar a modelagem do cenário do Turismo e analisar *frames* da experiência turística com os dados do português do Brasil, o escopo desta seção é tratar dos *corpora* compilados para o domínio do Turismo nesta língua. Para as demais línguas e domínios, trabalhos posteriores os especificarão.

O processo de criação de *corpus* demanda cuidados importantes. Sardinha (2004) pontua alguns pré-requisitos para que um conjunto de dados linguísticos possa ser considerado *corpus*. Esses dados, conforme elenca o autor, devem ser autênticos e em linguagem natural. Isso significa que precisam ser originários de falantes nativos e não podem ter sido criados visando à pesquisa linguística. Necessariamente, devem servir a um objeto de estudo e ser legíveis por computador. Acerca da composição, precisam ser representativos para o fim ao qual se destinam, sendo resultado de uma busca criteriosa.

Sobre o domínio do Turismo, respondendo aos critérios destacados em Sardinha (2004), os textos selecionados são autênticos, em linguagem natural e não foram criados para servir a uma pesquisa linguística. Vale ressaltar que um texto autêntico, na avaliação do autor, é aquele produzido por um falante nativo.¹ Quando esse critério é colocado em uma perspectiva multilíngue, *corpora* autênticos de mais de uma língua podem ser considerados comparáveis, diferentemente daqueles nomeados como paralelos, compostos por um texto na língua fonte e os demais sendo versões traduzidas. Quanto a esse último aspecto, o projeto inicial (SALOMÃO ET AL., 2011) previa a compilação de *corpora* paralelos e comparáveis, assim como realizado no Kicktionary (SCHMIDT, 2006, 2007, 2008, 2009), dicionário multilíngue para o domínio do futebol. O interesse era que as fontes de dados, sempre que possível, dispusessem das mesmas informações nas três línguas para que as entradas do dicionário fossem equivalentes. Com esse procedimento, a

¹ Quando os textos não forem produzidos por falantes nativos, o autor recomenda caracterizá-los como *corpora* de aprendizes – *learner corpora*. (SARDINHA, 2004, p. 19)

valência sintático-semântica representaria exatamente os equivalentes de cada sintagma e possibilitaria ainda contrastar esses padrões.

Todavia, esse propósito foi reconsiderado no que tange aos textos do Turismo (cf. GAMONAL, PERON & TORRENT, 2012). As traduções encontradas para a língua inglesa não condiziam com os padrões utilizados pelos falantes da língua. Algumas orações eram agramaticais, e alguns verbos não eram utilizados com o sentido almejado na língua alvo. Para o espanhol, os textos também tiveram problemas. Ainda que existissem cognatos com a língua portuguesa, essas palavras não eram empregadas para os mesmos fins. A avaliação foi que esses textos poderiam mesmo ter sido gerados por tradutores automáticos e não chegaram a ser tratados por especialistas.

Diante dessa realidade, seria um retrocesso utilizar tais textos, uma vez que muitos não poderiam ilustrar as entradas do dicionário e não possibilitariam que os padrões, tanto sintáticos quanto semânticos, representassem com segurança a realidade de cada um desses sistemas linguísticos. Por tais motivos, a alternativa mais coerente foi trabalhar com *corpora* comparáveis. Embora tenha sido abandonada a ideia de oferecer traduções das mesmas frases para ilustrar as entradas lexicais, o uso de texto comparável, retratando tipos de informações e fontes de dados semelhantes, corrobora o propósito de fazer do dicionário um guia de usos das línguas-alvo. É fato que os padrões de valência não serão comparados como se representassem a tradução de uma mesma sentença nas três línguas. Entretanto, diferentemente da análise de Bertoldi (2007), que avalia não ser possível mapear estruturas linguísticas em diferentes línguas se não forem utilizados *corpora* paralelos, o processo de anotação realizado por este projeto e pela FrameNet como um todo traz os padrões de combinação sintáticos e semânticos de todas as Unidades Lexicais² de um mesmo *frame*. Assim, o sumariamento desses padrões permite investigar as regularidades observadas em cada *frame*, possibilitando um contraste indireto entre línguas.

Além de explicitadas essas condições, Sardinha (2004) também elenca critérios que correspondem às principais nomenclaturas presentes na literatura da Linguística de *Corpus*. Seguem alguns: modo, tempo, seleção, conteúdo, autoria e finalidade. Atendendo a essa tipologia, os *corpora* coletados sobre turismo são:

² Adiante o conceito de Unidade Lexical será detalhado, por ora, cabe saber que uma Unidade Lexical é a união de uma forma linguística a um sentido específico.

escritos (modo); sincrônicos/contemporâneos, por designarem o período corrente, atual (tempo); dinâmicos, pois podem ser aumentados e/ou diminuídos (seleção); de domínios especializados (conteúdo); e, por último, são textos de falantes nativos (autoria).

Sobre a representatividade, apesar de ser muito discutida, não há um consenso na literatura que delimite regras para esse critério. Sardinha (2004) pontua que a primeira característica associada à representatividade é a extensão. Dessa forma, quanto maior o *corpus* mais representativo ele será. Porém, como ele mesmo pondera, perguntas como “representativo do quê?” e “representativo para quem?” são suscitadas nesse momento. Com tais questionamentos, o autor aponta a dificuldade de listar critérios que determinem a representatividade de um *corpus*. Ao avaliar essa questão, conseqüentemente, discute-se a adequação do *corpus*. Sardinha (2004) avalia que, por esse motivo, deve-se colocar a pesquisa à frente do objeto. Isso indica que os objetivos do estudo devem guiar as respostas sobre a representatividade e a adequação dos *corpora*.

Nesse sentido, os objetivos da presente pesquisa guiaram o processo de coleta do *corpus*. Pelo fato de o dicionário abordar a atividade turística no Brasil, país sede da Copa do Mundo de 2014, priorizou-se a busca por fontes de dados que tratassem dessa realidade. Os textos advindos de sites de órgãos governamentais de fomento à atividade turística – como aqueles organizados e apoiados pelo Ministério do Turismo brasileiro – e blogs de viajantes³ tiveram atenção especial. Primeiro, era preciso que as fontes de dados trouxessem informações seguras de quais atrações seriam encontradas em cada destino turístico bem como dados sobre a infraestrutura desses locais. Para isso, os textos organizados ou apoiados pelo governo foram eficazes, além disso, alguns desses materiais foram atualizados para incentivar a atividade turística durante o evento da Copa. Segundo, como se trata de uma atividade essencialmente humana, era interessante explorar as experiências dos turistas acerca de cada lugar visitado.

³ Os blogs que compõem os *corpora* desta pesquisa fazem parte da RBBV (Rede Brasileira de Blogueiros de Viagem). Fonte: <<http://www.rbbv.com.br/>>

<i>Corpus</i>	<i>URL</i>	<i>Descrição</i>	<i>Tokens</i>
Brazil_Tour	www.braziltour.com	Organizado pelo Instituto Brasileiro de Turismo (EMBRATUR). Esse site de fomento à atividade turística no Brasil trata de variados locais que o turista pode optar caso tenha interesse em alguma atividade turística específica, como atrações buscadas com foco no aspecto cultural, ecoturismo, sol e praias, turismo focado nos negócios e no esporte.	51.164
Brasil_Gov	www.brasil.gov.br	Página do governo federal que trata de variados assuntos, entre eles o turismo, abordando dicas para quem pretende visitar o país, precauções que se devem tomar, tipos de turismo, principais atrações para quem deseja algo específico, como festas regionais ou negócios, por exemplo, notícias recentes sobre viagens, infraestrutura etc.	18.623
Vai_Brasil	www.vaibrasil.com.br	Projeto criado pelo Ministério do Turismo em parceria com a Associação Brasileira das Operadoras de Turismo (BRAZTOA) e a Associação Brasileira das Agências de Viagem (ABAV) com o intuito de fomentar a comercialização de pacotes turísticos nos diversos destinos pelo Brasil.	91.483
Comi_erninha_de_cachorro	www.comiperninhade cachorro.com	O blog nasceu em 2009, quando os autores decidiram fazer um intercâmbio pela Austrália e daí cresceu e está se tornando um apanhado geral de viagens e experiências.	17.382
Destino_de_Viagens	www.destinodeviagem.com.br	O blog reúne experiências em vários países e aventuras pelo mundo. Atualmente, busca novos destinos e estrutura roteiros turísticos pelo Brasil.	18.918
Andarilhos_do_mundo	andarilhosdomundo.com.br	O blog é um projeto de dois amigos que decidiram compartilhar as aventuras em suas viagens.	4.891
Total de tokens			202.461

Quadro 1: *Corpora* compilados para o domínio do Turismo

A observação dessas características foi fator determinante para a seleção dos *corpora*, que apresentam o conteúdo esperado e marcações tanto da variante escrita padrão quanto não padrão da língua. Sobre sua extensão, se se considerarem os estudos realizados por Sardinha (2004), a avaliação é que se trata de *corpus* pequeno médio (80 a 250 mil palavras), pois, de acordo com o explicitado no Quadro 1, os *corpora* apresentam um total de 202.461 *tokens* / palavras.

Além dos *corpora* compilados para representar o evento turístico, outro também constituído pela FrameNet Brasil é o FIFA.⁴ Os textos que estão sendo utilizados para apoiar a estruturação do Cenário da Copa advêm do site da Federação Internacional de Futebol, www.fifa.com, e são disponibilizados em várias línguas, dentre elas o Português, Inglês e Espanhol. O *corpus* compilado para a língua portuguesa contém mais de 400 mil *tokens*, distribuídos em diversas notícias de futebol e eventos anteriores da Copa. Como também estão inclusas informações turísticas do Brasil, mas principalmente sobre a última Copa do Mundo, realizada na África do Sul, sentenças desse banco de dados são utilizadas por este trabalho, ainda que não, necessariamente, exemplifiquem as entradas do dicionário.

A compilação dos *corpora* do Turismo se deu manualmente. Os autores dos blogs, em grande parte, escrevem diretamente no corpo do site antes de publicar os textos, o que dificultou o envio dos documentos. O mesmo trabalho foi realizado para os textos disponibilizados por órgãos do Ministério do Turismo, que autorizou o processo de coleta da mesma forma. Sobre o armazenamento, esses dados foram reunidos em arquivos em formato *.txt* para que fossem devidamente reconhecidos pelo PALAVRAS (BICK, 2000), *parser* que etiqueta os dados morfossintaticamente.

Não houve uma regra específica para organizar os arquivos. Como a quantidade de sites foi pequena, convencionou-se identificar os documentos com o nome do *corpus*. Quando foi necessário dividir em mais de um arquivo, foi destacado o nome do *corpus* seguido da região ou temática, como sol e praia. A Figura 1 ilustra um *corpus* em arquivo *.txt*. A marcação explicitada no corpo do texto é necessária para sua identificação no *software* FrameNet Desktop e também para fornecer em *link* para o usuário a fonte de cada sentença.

⁴ A análise realizada para os textos encontrados em sites multilíngues do domínio do Turismo não é estendida para os textos disponibilizados pelo site da FIFA, pois estes são versões traduzidas que representam a realidade da língua de cada país.

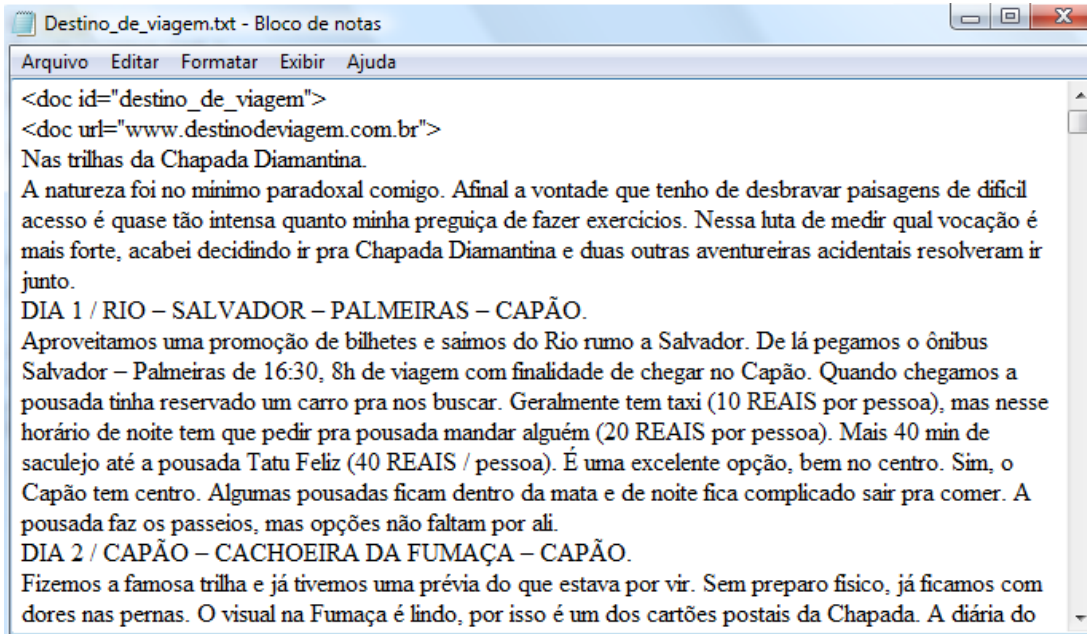


Figura 1: Texto em formato *.txt*

1.1.2 Tratamento Computacional

Concluído o pré-processamento, que determina os critérios de escolha, coleta e procedimentos necessários para o armazenamento e processamento dos textos, parte-se para o tratamento computacional, propriamente dito. Nessa fase, são realizadas diversas tarefas que etiquetam os textos com análises e marcações necessárias para que, posteriormente, sejam segmentados.

A ferramenta escolhida para realizar essas tarefas foi o *parser* PALAVRAS (BICK, 2000). Esse *software* é um analisador automático desenvolvido para o português por Eckhard Bick. Ele apresenta um léxico com milhares de lemas e regras gramaticais⁵, que fornecem análise morfológica e sintática aos textos. Sua atuação é dividida em módulos. A primeira etapa é realizada pelo módulo PALTAG. Nele, o analisador PALMORF cria as fronteiras necessárias entre palavras e sentenças para que o texto seja analisado. É esse processador que estabelece o primeiro contato com o texto. Nesse módulo, são reconhecidas e realizadas as possíveis leituras para cada palavra, identificando-se abreviações, lexemas, palavras compostas, expressões polilexicais, flexões, nomes próprios, dentre vários

⁵ As regras gramaticais realizadas por este *parser* são baseadas na Gramática das Restrições – *Constraint Grammar*. Após o texto ter sido previamente tratado, essas regras especificam as leituras que podem e não podem ser realizadas em cada palavra (cf. BICK, 2000).

outros fenômenos. Após essas informações serem identificadas, elas são submetidas às regras da Gramática das Restrições, que processarão esses dados e buscarão soluções para as possíveis ambiguidades. Nesse momento, passa-se para a atuação do próximo módulo, o PALSYN, que faz o mapeamento sintático, atribuindo às classes de palavras ou aos lemas as funções sintáticas. Nessa etapa, os dados são filtrados e ainda são sugeridas possíveis soluções para a resolução de ambiguidades. Além desses módulos, o PALAVRAS inclui outros, o PALSEM e o PALTRANS. O primeiro é um módulo responsável por propor soluções para as ambiguidades que persistirem, e o segundo trata de traduções para outras línguas.

O *software* é capaz de gerar saídas a partir de cada um dos módulos acima. Como o objetivo do processamento dos textos pelo PALAVRAS é possibilitar a busca por lexemas específicos nos vários contextos sintáticos em que aparecem, utiliza-se, para o Copa 2014, a saída do arquivo fornecida pelo módulo PALSYN. O processamento dos *corpora* é realizado via acesso remoto ao servidor do grupo SEMANTEC, liderado pela Profa. Rove Chishman na UNISINOS e parceiro da FrameNet Brasil no desenvolvimento do dicionário.

A saída do PALSYN (no PALAVRAS) é submetida a outro tratamento computacional, através da ferramenta SYN-to-TAG, desenvolvida pela FrameNet Brasil. Tal ferramenta promove uma série de alterações no arquivo .syn a fim de criar um arquivo .tag adaptado para ser compatível com a ferramenta *Word Sketch*. Entre estas alterações estão a retirada de caracteres especiais usados nas pontuações e as adaptações das contrações (por exemplo, de+o = do). A próxima etapa do tratamento computacional, conforme ilustrado pela Figura 2, é realizada pelo Sketch Engine (KILGARRIFF, RYCHLY, SMRZ & TUGWELL, 2004).

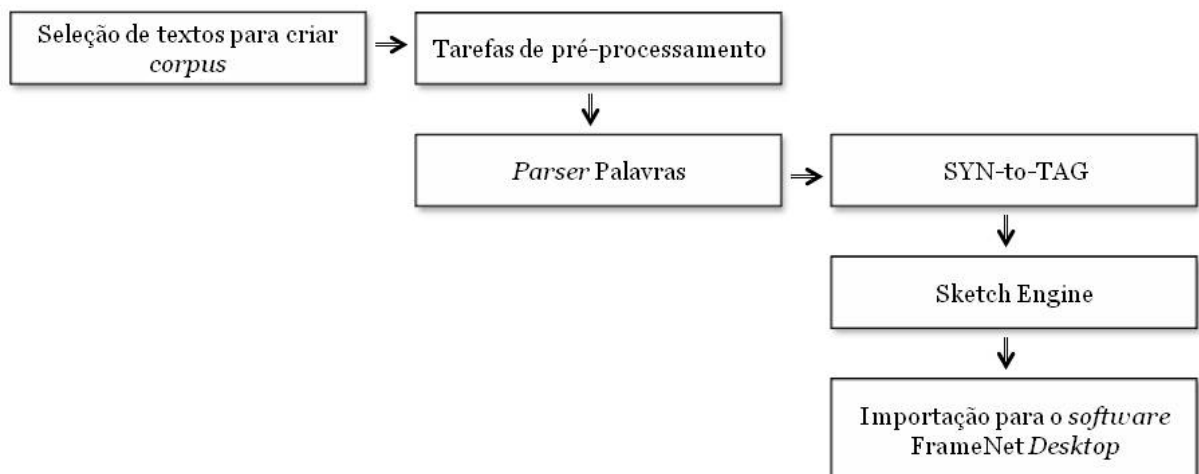


Figura 2: Etapas do tratamento computacional

O Sketch Engine, disponível em <<http://www.sketchengine.co.uk>>, é um sistema de consulta a *corpus*. Com ele é possível utilizar *corpora* de grande extensão, além de poder abrigar aqueles criados pelo usuário. A FrameNet Brasil utiliza essa ferramenta para hospedar os *corpora* compilados pelo projeto e realizar o processo de importação desses documentos para o FrameNet Desktop, *software* no qual o procedimento de anotação dos dados é realizado. Para a importação das sentenças, dois recursos dessa ferramenta podem ser utilizados, o *Concordance* e o *Word Sketch*. A busca pelo *Concordance* retorna todos os dados encontrados nos *corpora*. Na Figura 3, a pesquisa foi feita por lema, e o buscado foi *turista*. Juntamente com as sentenças, fornecem-se as fontes de cada *corpus*.

Corpus: COPA 2014 FN-Br	
Hits: 297 (476.7 per million)	
First Previous Page <input type="text" value="5"/> of 6 Go Next Last	
Portal Brasil - Turismo, http://www.brasil.gov/sobre/turismo	é registrado no Sistema=Câmbio . É obrigatório apresentar comprovantes de como o turista conseguiu ac
Portal Brasil - Turismo, http://www.brasil.gov/sobre/turismo	atividades ilícitas como lavagem de dinheiro . Para avaliar qual é a melhor cotação , o turista deve compar
Comi perninha de cachorro pelo mundo, http://www.comiperninhadecachorro.com	grande janela para o rio , introduzindo essa cultura ao povo de Belém e atraindo mais turistas . O projeto t
Comi perninha de cachorro pelo mundo, http://www.comiperninhadecachorro.com	Canoa=Quebrada, difícil é comer mal , lagostas e camarões a preços ótimos , para turista nenhum bota
Comi perninha de cachorro pelo mundo, http://www.comiperninhadecachorro.com	ambos instrutores de escalada da Multisport , estiveram lado=a=lado ajudando nos turistas leigos a reali
Comi perninha de cachorro pelo mundo, http://www.comiperninhadecachorro.com	aluguel , enfileirados e apertados nas suas baías eles esperam por ali para levar os turistas para passear
Destino de Viagem, www.destinodeviagem.com.br	não me deixou uma impressão espetacular , a cidade esta mal conservada e vi muitos turistas comentando
Destino de Viagem, www.destinodeviagem.com.br	extraordinário e de ter visto o deserto na sua forma mais selvagem , bem longe da multidão de turistas que lotam os
Destino de Viagem, www.destinodeviagem.com.br	lado e é bem mais isolado e autêntico . na alta temporada Caburé é invadida por turistas com quadric
Destino de Viagem, www.destinodeviagem.com.br	, praias isoladas e limpinhas , paraísos tropicais do nordeste . Jeri tem muito turista , tem agito ,
Destino de Viagem, www.destinodeviagem.com.br	Transpantaneira não são fazendas na sua maioria , mas pousadas construídas pra receber turistas e fotógrafos
Destino de Viagem, www.destinodeviagem.com.br	não fosse trágico . O que é a hotelaria carioca ? Em=vez=de honras e pompas , os turistas geralmente e
Destino de Viagem, www.destinodeviagem.com.br	as melhores opções pra um publico que leva a=sério a relação custo \$/ benefício o turista francês . O n
Destino de Viagem, www.destinodeviagem.com.br	catástrofes (que são muitas) e a lista vai ser alimentada constantemente . Copacabana . Turista que vem pela
BrazilTour Ecoturismo, http://www.braziltour.com	rios de águas cristalinas repletos de peixes , suas grutas e cachoeiras que atraem turistas do mundo to
BrazilTour Ecoturismo, http://www.braziltour.com	pela perfeita harmonia entre o mar e a montanha . Suas inúmeras atrações trazem turistas de todo=om
BrazilTour Ecoturismo, http://www.braziltour.com	a atriz Brigitte=Bardot se encantou pelo local atraindo uma grande quantidade de turistas . Atualmente
BrazilTour Ecoturismo, http://www.braziltour.com	do entorno que a complementa . As suas manifestações artístico-culturais atraem turistas de todos=os
BrazilTour Cultura, http://www.braziltour.com	anda a=pé ou de charrete . E é o lugar aonde a prática do kitesurf atrai vários turistas europeus . E
BrazilTour Cultura, http://www.braziltour.com	roteiro é possível avistarem-se e visitar as salinas artesanais e industriais . O turista poderá enco
BrazilTour Cultura, http://www.braziltour.com	Cerro=Corá , Currais=Novos , Jardim=do=Seridó , Parelhas . Um roteiro diferente onde o turista parte na dire
BrazilTour Cultura, http://www.braziltour.com	Porto=de=Galinhas e Muro=Alto são verdadeiros cartões-postais , de deixar qualquer turista encantado .
BrazilTour Cultura, http://www.braziltour.com	do Brasil . E especialmente , de todo=o litoral nordestino . É neste local que os turistas encontram a
BrazilTour Cultura, http://www.braziltour.com	possível ver parte da significativa das pinturas=rupestres . O parque é visitado por turistas e estudiosos

Figura 3: Resultado da busca pelo lema *turista* na ferramenta *Concordance*

Fonte: <<https://the.sketchengine.co.uk>>

O *Word Sketch* oferece ao usuário uma busca mais refinada. Para o lema *turista*, conforme mostra a Figura 4, são fornecidos os contextos sintáticos em que essa palavra é encontrada. A utilização desse recurso para os dados na língua portuguesa requer que os *corpora* tenham sido anteriormente filtrados pelo *parser* PALAVRAS, responsável pelo tratamento morfossintático dos textos. Essa ferramenta trabalha por amostragem, fazendo com que as combinações sintáticas mais recorrentes sejam visualizadas.

Concordance		Word List		Word Sketch		Thesaurus		Find X		Sketch-Diff		?		
<p>turista COPA 2014 FN-Br freq = 198 (380.3 per million)</p>														
modificar 26 1.2			sujeito de 43 4.1			objeto de 25 1.9			pp de+o 7 0.8			pp para 2 1.0		
encantado 3 11.73			partir 4 10.73			atrair 10 11.06			mundo 6 8.9			primórdio 1 13.0		
estrangeiro 6 11.0			aproveitar 7 10.16			deslumbrar 1 10.25			país 1 5.25			cidade 1 4.73		
leigo 1 10.25			atrair 4 9.57			transportar 1 9.09								
europeu 4 9.45			entrevistar 1 9.51			trazer 3 9.3			pp de 6 0.6			pp durante 1 6.0		
nacional 3 8.99			encantar 2 9.43			chamar 1 8.45			cruzeiro 1 10.91			campeonato 1 9.48		
brasileiro 7 8.88			gastar 1 9.42			deixar 3 8.23			negócio 3 9.26					
francês 1 8.66			abastecer 1 9.36			levar 2 7.86			canto 1 9.02			pp em+o 1 0.2		
prático 1 8.57			encontrar 6 9.14			receber 2 6.76			estado 1 7.42			local 1 5.98		
			possibilitar 1 8.93			ter 2 4.15								
			permanecer 1 8.85						pp com 6 2.3					
			comentar 1 8.81						quadríciclo 1 12.19					
			percorrer 1 8.77			e ou 22 1.7			olho 2 10.39					
			precisar 2 8.58			estudioso 3 11.94			paisagem 2 9.13					
			visitar 2 8.25			morador 5 11.17			opção 1 6.78					
			integrar 1 8.22			fotógrafo 1 10.05								
			compor 1 8.18			executivo 4 10.0			pp por 3 4.1					
			querer 1 7.78			velejador 1 9.96			mês 2 8.84					
			vir 1 7.48			proprietário 1 9.68			ano 1 5.69					
			garantir 1 7.06			instituição 1 9.45								
			conseguir 1 6.85			câmera 1 9.12			pp em 2 0.8					
						visitante 2 8.65			escolha 1 8.79					

Figura 4: Resultado da busca por *turista* na ferramenta *Word Sketch*

Fonte: <<https://the.sketchengine.co.uk>>

Veja que, quando *turista* assume função de sujeito, Figura 4, são especificadas as ações verbais encontradas com esse sintagma, por exemplo, *aproveitar*, *entrevistar* e *visitar*. Para fins do dicionário, o mais provável é que *entrevistar* não faça parte da busca, pois, geralmente, não atua no domínio turístico, mas *aproveitar* e *visitar*, sim. Se a pesquisa for realizada com *turista* assumindo função de objeto, sentenças cujos verbos são *atrair*, *deslumbrar* e *trazer* apresentam grandes chances de retratar um contexto turístico. Com a frequência retornada para cada um desses sentidos, é possível especificar qual agrupamento de sentenças será importado para o FrameNet Desktop.

A importação dos textos pode ser realizada tanto via *Concordance* como via *Word Sketch*. Como os *corpora* são de domínio especializado, o mais natural é que os lemas buscados representem a Unidade Lexical desejada. Por exemplo, o lema *visitar* provavelmente tratará de visitas a atrações turísticas, assim a importação pelo *Concordance* é a mais indicada, pois todos os dados são importados, evitando que sentenças sejam perdidas pela amostragem do *Word Sketch*. Todavia, há casos em que o mais indicado é uma busca mais refinada. Por exemplo, no *frame* *Atração_em_lugar*, a Unidade Lexical *ter* o evoca em sentenças do tipo: *Canela tem cenários naturais belíssimos – corpus Vai_brasil*. Como esse lema é retornado em grande quantidade, com dados que não fazem parte do *frame*, o ideal é que a

importação aconteça via *Word Sketch* para especificar os tipos de sintagmas que acompanham esse lema. O refinamento possibilitado pelo *Word Sketch* especifica quais são as funções sintáticas assumidas pelo lema buscado e exhibe vários sintagmas que o acompanham, tornando possível conhecer os contextos de atuação e, conseqüentemente, de quais *frames* o lema participa.

Para os dois procedimentos de importação, algumas medidas são necessárias, como especificar a extensão do contexto de cada sentença para que os dados sejam compreendidos – a esse procedimento, convencionou-se uma extensão de 1000 caracteres. A ferramenta de importação traz para o Desktop as sentenças segmentadas pela existência de um sinal de pontuação (., ! e ?) antes e um depois do lema pesquisado. Nesse procedimento, pode ser que se percam algumas sentenças muito extensas.

Depois de realizado o processo de importação das sentenças, o FrameNet Desktop (FN Desktop) garante a análise dos dados nos padrões da FrameNet mãe (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010). A Figura 5 ilustra como o *corpus* aparece nesse *software*, que destaca as Unidades Lexicais de preto e em caixa alta, vide a UL *turismo*. O lexema marcado pode ser Unidade Lexical de outro *frame*, quando isso acontece, a sentença com esse lexema não é anotada. Na Figura 5, esse fenômeno não acontece, pois o lexema *turismo*, necessariamente, evocará o *frame* *Cenário_do_turismo*.

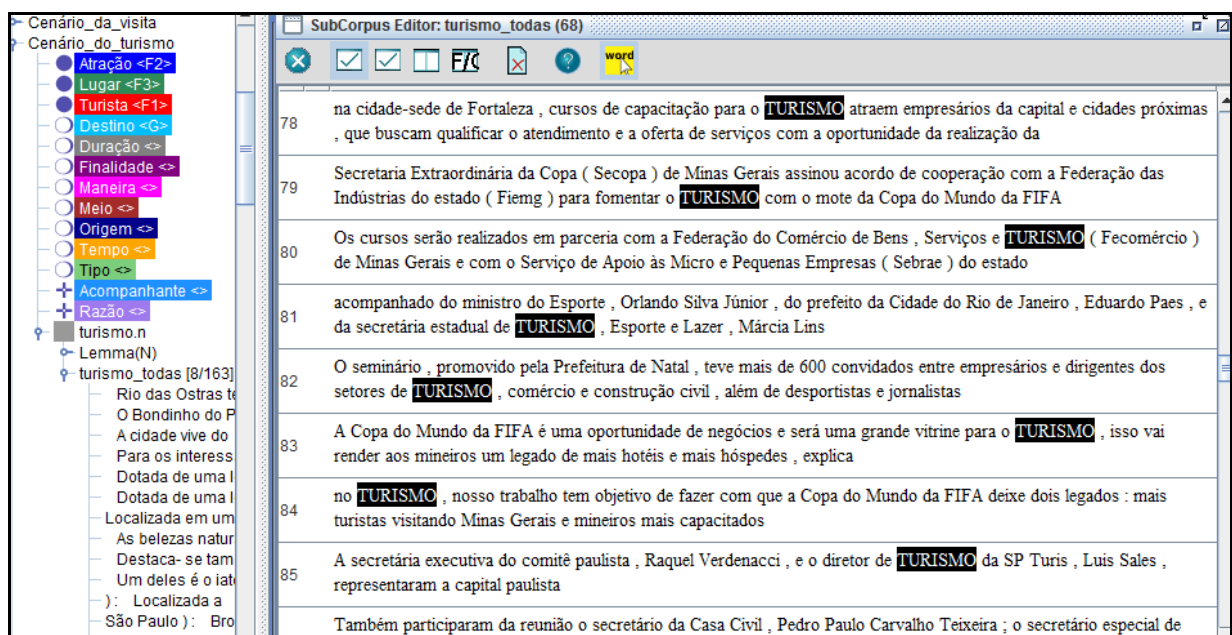


Figura 5: *Corpora* no *software* FrameNet Desktop

Antes de as sentenças serem analisadas, procedimento que será detalhado no desenvolver deste trabalho, os *frames* bem como seus elementos e Unidades Lexicais são definidos nessa plataforma.

1.2 Métodos para Estruturação de Frames

Na estruturação de *frames*, duas metodologias podem ser seguidas: *bottom-up* e *top-down*. No método *bottom-up*, o anotador parte dos dados para criar o *frame*. A primeira tarefa é selecionar, intuitivamente, um agrupamento de lexemas que se relacionam semanticamente, por exemplo, *visitar*, *turista*, *apreciar*, *visitante* e *atração*. Estes lexemas, ainda que apresentem particularidades quanto ao sentido e não pertençam todos às mesmas classes de palavras, participam de um mesmo domínio semântico, aquele que trata de atividades turísticas.

Em um segundo momento, parte-se para a pesquisa dos itens lexicais selecionados nos *corpora*, tendo em mente um possível *frame* para investigação. Após esse estudo, selecionam-se algumas sentenças com esses itens lexicais com o objetivo de analisar o comportamento desses predicadores no que tange à valência sintático-semântica. Com isso em mãos, o analista procura regularidades tanto semânticas quanto sintáticas que permitam a estruturação de uma situação específica, essa etapa é a definição do *frame*. Nela, são determinadas as Unidades Lexicais (ULs), previamente selecionadas pelo analista, são especificadas a nuclearidade dos Elementos de *Frame* (EFs) bem como a relação entre os EFs. Quando o *frame* estiver estruturado, é possível relacioná-lo com outros, caso exista entre eles alguma relação, essa é a etapa de assinalar as relações entre *frames*.

No método *top-down*, as primeiras análises são delineadas pela intuição de falante nativo ou conhecedor da língua, o anotador propõe uma definição para o *frame*, elencando as possíveis Unidades Lexicais que o evocam, designando a nuclearidade de cada elemento, traçando as possíveis relações entre *frames* e EFs. Em um segundo momento, verifica-se como foi o seu comportamento nos *corpora* para confirmar se essa definição condiz com os dados. Se necessárias, alterações na definição do *frame* e dos EFs podem ser realizadas.

Geralmente, considera-se que o método *bottom-up*, por partir, primeiramente, dos dados, possibilita uma análise mais ancorada na realidade, sendo mais fiel aos

dados que o método *top-down*, que cria o *frame* para depois investigá-lo no *corpus*. Entretanto, como ambos os métodos são submetidos à apreciação em *corpus*, permitindo que os dados encontrados definam os *frames* e, conseqüentemente, as considerações acerca de seus elementos, esses métodos não se opõem e, por isso, não devem ser comparados com teor de inferioridade e/ou superioridade.

Para construir os *frames* do Turismo, num primeiro momento, foi priorizado o método *bottom-up*, que partia do mais básico, as Unidades Lexicais, para aquilo mais abstrato, os *frames*. Num segundo momento, o método adotado foi *top-down*.

1.3 Anotação Lexicográfica na FrameNet Brasil

A FrameNet Brasil, assim como a FrameNet mãe, prevê dois tipos de anotação de dados: texto corrido e lexicográfico. A anotação de texto corrido trata de todos os *frames* que forem surgindo no texto escolhido para análise. Nesse procedimento, o texto é o responsável por guiar o aparecimento das ULs e, conseqüentemente, a análise dos *frames*. Já, na anotação lexicográfica, a pesquisa acontece a partir de uma Unidade Lexical específica.

Para a elaboração do dicionário, o método seguido foi o lexicográfico. Como o objetivo de um dicionário de domínio especializado é tratar de um determinado conjunto de palavras, a anotação lexicográfica é a mais indicada. Com esse procedimento, é possível verificar a atuação das Unidades Lexicais em diferentes materiais linguísticos.

As etapas de análise de dados são as mesmas tanto para a anotação lexicográfica quanto para a de texto corrido. Esses procedimentos são devidamente aprofundados nas linhas que se seguem neste trabalho. De modo geral, cabe saber que o processo de anotação da FrameNet é realizado, majoritariamente, em três camadas, Elemento de *Frame* (no inglês FE, *Frame Element*), Função Gramatical (no inglês GF, *Gramatical Function*), e Tipo Sintagmático (no inglês PT, *Phrase Type*). Como as palavras que evocam *frames* são predicadores, como substantivos, adjetivos e verbos, especifica-se a valência desses itens lexicais tanto em relação à sintaxe quanto à semântica.

The screenshot shows the FrameNet Brasil interface. At the top, a sentence is displayed: "O Jalapão apresenta um panorama de perder o fôlego". Below the sentence is a grid of annotations. The first row, labeled "Layer", contains the words of the sentence. The second row, "FE", shows "Lugar" (green) under "Jalapão" and "Atração" (blue) under "um panorama de perder o fôlego". The third row, "GF_br", shows "Ext" (red) under "Jalapão" and "ObjD" (blue) under "um panorama de perder o fôlego". The fourth row, "PT_br", shows "NP" (green) under "Jalapão" and "NP" (blue) under "um panorama de perder o fôlego". Below the grid is a list of frame elements for the "Atração" frame:

<input checked="" type="radio"/> Atração <F4>		<input type="radio"/> Meio <>	
<input checked="" type="radio"/> Lugar <F11>		<input type="radio"/> Origem <>	
<input checked="" type="radio"/> Turista <F1>	DNI	<input type="radio"/> Tempo <F2>	
<input type="radio"/> Descrição <>		<input type="radio"/> Acompanhante <>	
<input type="radio"/> Duração <>		<input type="radio"/> Frequência <>	
<input type="radio"/> Finalidade <>		<input type="radio"/> Razão <F8>	
<input type="radio"/> Maneira <>			

Figura 6: Processo de anotação em três camadas na FrameNet Brasil

Na Figura 6, a Unidade Lexical *apresentar* na sentença *O Jalapão apresenta um panorama de perder o fôlego* evoca o frame *Atração_turística*. Sua valência é preenchida pelo FE Lugar, *O Jalapão*, com função gramatical Externo e tipo sintagmático Sintagma Nominal, e *um panorama de perder o fôlego*, EF Atração, com FG Objeto Direto e TS Sintagma Nominal. O EF Turista não é instanciado na sentença, mas atua como Instanciação Nula Definida, pois, ainda que não seja retomado pelo texto, infere-se que o turista seja o ator principal de todo esse sistema. Um argumento para isso é o fato de todos os *corpora* compilados para este trabalho retratarem contextos turísticos.

2 RECURSOS LEXICAIS ELETRÔNICOS PARA USO HUMANO E PARA APRENDIZADO DE MÁQUINAS

Estruturar um dicionário requer, além do material lexical, um suporte teórico capaz de responder aos diversos fenômenos linguísticos com os quais o lexicógrafo cotidianamente se depara. Para Borba (2003), a lexicografia pode ser vista como técnica e teoria: quando analisada como técnica, o foco se dirige à estruturação de verbetes, à seleção das entradas, ao registro de variantes etc; se analisada como teoria, são levados em conta os princípios que possibilitam a descrição do léxico, permitindo, assim, uma metalinguagem capaz de manipular e apresentar as informações pertinentes. Nesse sentido, a escolha de uma teoria linguística é crucial para amparar as respostas para as diversas relações estabelecidas no léxico.

Ainda que avaliando que esses dois aspectos estejam intimamente relacionados, de modo que a opção por um interfira direta ou indiretamente na estruturação do outro, neste trabalho, o empenho maior será mostrar como a teoria linguística da Semântica de *Frames* (FILLMORE, 1982, 1985) somada à metodologia da FrameNet (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010) podem contribuir na montagem de dicionários eletrônicos multilíngues. O tratamento mais específico relativo às escolhas técnicas e o consequente refinamento do *software* utilizado serão aprofundados em trabalhos posteriores.

Para cumprir tais propósitos, a organização deste capítulo se dispõe da seguinte forma: na seção 2.1, abordam-se os dicionários eletrônicos, com o intuito de contrastá-los com os impressos e expor os desafios colocados na confecção de versões eletrônicas multilíngues; nas seções 2.2 e 2.3, apresentam-se a FrameNet e a WordNet, dois recursos lexicais estruturados eletronicamente; na seção 2.4, o Kicktionary é ilustrado como proposta concluída de dicionário multilíngue de domínio específico que utiliza *frames* e *synsets*.

2.1 Dicionários Eletrônicos

Os dicionários veiculados em meio eletrônico começaram a ser amplamente discutidos pela comunidade acadêmica e produzidos em larga escala a partir da década de noventa, quando Schryver (2003) avalia que houve seu grande *boom*. As discussões dos lexicógrafos acerca de como seria o “dicionário do futuro” bem como

quais seriam os usuários e os lexicógrafos do futuro passaram a se tornar então palpáveis.

Quando se fala em dicionário eletrônico, é importante ter em mente a quem o recurso é direcionado, se a usuários humanos ou a máquinas. Isso porque, além de uso humano, eles também servem para aprendizado de máquina, ou seja, para tarefas de Processamento de Linguagem Natural (PLN). Há recursos que são legíveis para os dois públicos, como é o caso da FrameNet (FILLMORE ET AL., 2003; RUPPENHOFER ET AL., 2010) e da WordNet (FELLBAUM, 1998; MILLER ET AL., 1993). Uma definição que contempla essa realidade é dada por Nesi (2000 apud SCHRYVER, 2003):⁶

O termo dicionário eletrônico (ou ED) pode ser usado para se referir a qualquer material de referência armazenado em formato eletrônico que fornece informações sobre ortografia, significado ou uso de palavras. Assim, um corretor ortográfico em um programa de processamento de texto, um dispositivo que analisa e traduz palavras impressas, um glossário para materiais de ensino *on-line*, ou uma versão eletrônica de um dicionário em papel são todos tipos de dicionários eletrônicos. (NESI 2000 apud SCHRYVER, 2003 p. 145)⁷

Adotando uma definição ampla, que agrupe diferentes recursos lexicais, Schryver (2003) propõe uma tipologia que classifique um dicionário eletrônico sem negligenciar pontos relevantes acerca de sua constituição. Desse modo, considera como prioridade definir sua forma de acesso. Com esse prisma, o autor pontua que se deve ter em mente a seguinte pergunta: “Quem acessa o que onde?”. Se a análise tem início a partir do *quem*, é possível que seja um usuário humano ou uma máquina; se for humano, o léxico deve ser legível para humano, caso seja máquina, precisa ser compatível com tarefas de PLN. A respeito da pergunta *o que*, em linhas gerais, procura-se definir se o recurso é ou não eletrônico. Já a última pergunta, *onde*, analisa os dispositivos de armazenamento para hospedar tais dicionários. Mantendo o foco nas versões eletrônicas, os dispositivos de armazenamento podem ser *offline* ou subsidiados pela *web*. Conforme a Figura 7 ilustra, os dicionários eletrônicos *offline* podem ser desenvolvidos para atuarem em máquinas de computador ou como versões *pocket*, com opção de adotar uma aparência

⁶ Todas as traduções das obras em língua estrangeira citadas neste trabalho, salvo menção em contrário, são de responsabilidade da autora. As citações traduzidas serão acompanhadas da versão original, reproduzida em nota de rodapé.

⁷ “The term electronic dictionary (or ED) can be used to refer to any reference material stored in electronic form that gives information about spelling, meaning, or use of words. Thus a spell-checker in a word-processing program, a device that scans and translates printed words, a glossary for on-line teaching materials, or an electronic version of a respected hard-copy dictionary are all EDs of a sort”.

inovadora ou seguir o modelo proposto pelos tradicionais de papel. Semelhantes são os recursos que atuam em ambiente *online*, pois, ainda que sejam amparados pela *web*, podem também seguir um modelo tradicional ou optar por uma versão autêntica.

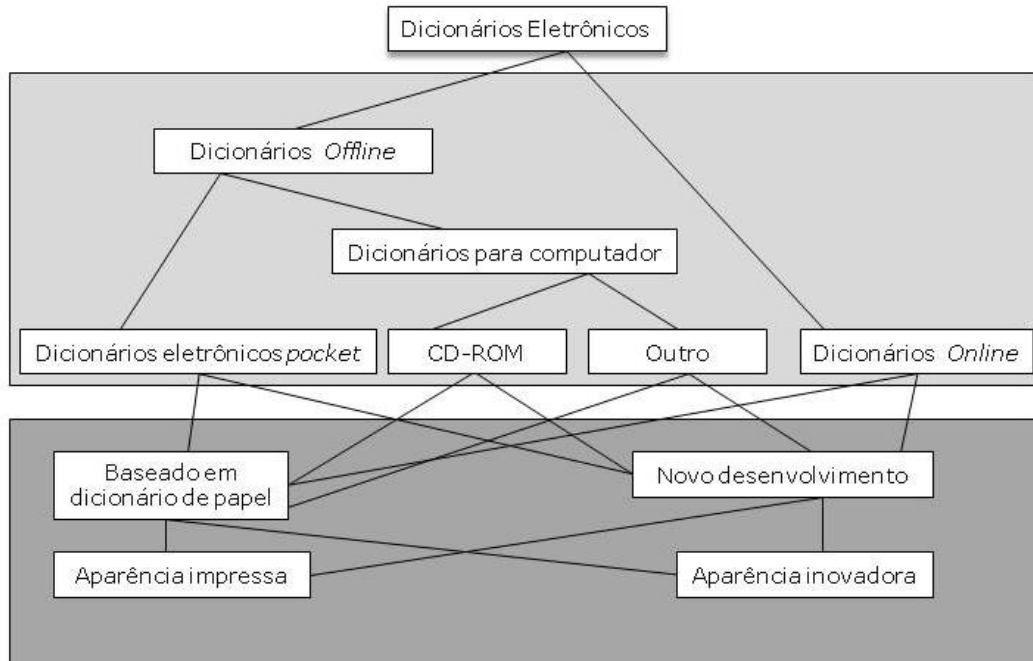
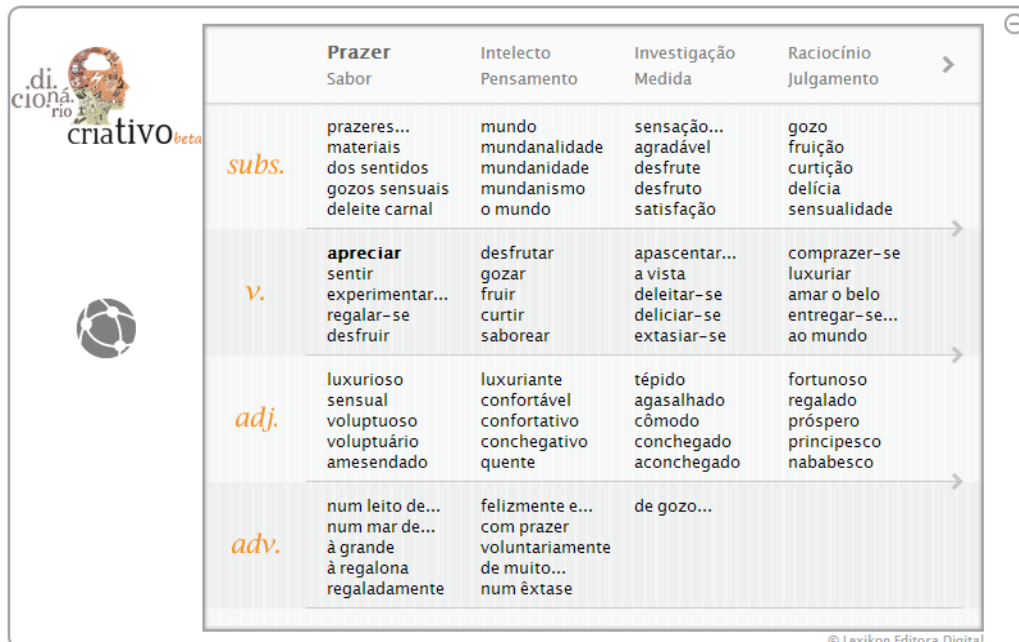


Figura 7: Dicionário eletrônico *offline* e *online*
 Fonte: Adaptado de Lehr (1996 apud SCHRYVER, 2003, p 146)

O dicionário Copa 2014, conforme guiado pela Figura 7, segue o caminho dos dicionários eletrônicos de versão *online* com interface inovadora e desenvolvimento diferente daquele adotado nas versões impressas. Em consonância com Borba (2003), que avalia que os dicionários não podem ser tomados apenas como simples repositórios ou acervos de palavras, mas como guias de uso, a opção por organizar os dados por *frames* reconhece que dicionários são estruturados tanto com informações linguísticas como com informações enciclopédicas. Desse modo, como o sistema de busca será guiado por *frames*, as palavras estarão agrupadas por campos semânticos, e as entradas conterão dados a respeito dos elementos participantes dos *frames*. Essa gama de informações auxiliará na estruturação das entradas, evitando que as definições sejam tão falhas como acontecem em muitos momentos com os dicionários tradicionais.

Os dicionários analógicos ou *thesauri* são recursos que se aproximam desse propósito. O fato de não utilizarem uma definição nos verbetes faz com que o agrupamento de palavras pertencentes a um mesmo campo semântico cumpra esse

papel. Um recurso eletrônico *online* interessante para o português brasileiro é o Dicionário Criativo, www.dicionariocriativo.com.br, atualmente disponível na versão beta.



The screenshot shows a search result for the word 'apreciar' in the 'Dicionário Criativo' interface. The interface is organized into a grid with four columns representing different semantic fields: 'Prazer', 'Intelecto', 'Investigação', and 'Raciocínio'. The search results are categorized by part of speech: 'subs.' (substantivo), 'v.' (verbo), 'adj.' (adjetivo), and 'adv.' (advérbio). Each category lists related words and phrases. The interface also includes a logo for 'dicionário criativo beta' and a copyright notice for 'Lexikon Editora Digital'.

	Prazer Sabor	Intelecto Pensamento	Investigação Medida	Raciocínio Julgamento
subs.	prazeres... materiais dos sentidos gozos sensuais deleite carnal	mundo mundanidade mundanidade mundanismo o mundo	sensação... agradável desfrute desfruto satisfação	gozo fruição curtição delícia sensualidade
v.	apreciar sentir experimentar... regalar-se desfruir	desfrutar gozar fruir curtir saborear	apascentar... a vista deleitar-se deliciar-se extasiar-se	comprazer-se luxuriar amar o belo entregar-se... ao mundo
adj.	luxurioso sensual voluptuoso voluptuário amesendado	luxuriante confortável confortativo conchegativo quente	tépido agasalhado cômodo conchegado aconchegado	fortunoso regalado próspero príncipesco nababesco
adv.	num leito de... num mar de... à grande à regalona regaladamente	felizmente e... com prazer voluntariamente de muito... num êxtase	de gozo...	

Figura 8: Busca por *apreciar* na seção do dicionário analógico no Dicionário Criativo
Fonte: <<http://www.dicionariocriativo.com.br>>

Essa iniciativa de Felipe Iszlaji de Albuquerque⁸ reúne diferentes ferramentas eletrônicas disponíveis *online* com o objetivo de agregar em um único recurso um material lexicalmente rico e ao mesmo tempo eficiente para o usuário. A Figura 8 ilustra, com o lema *prazer*, o resultado da seção que disponibiliza as informações próprias a um dicionário analógico. As palavras são agrupadas pelas categorias lexicais às quais pertencem: por exemplo, na categoria *substantivo*, são encontradas palavras como *satisfação*, *gozo*; e, na categoria *verbo*, *desfrutar*, *saborear*. O usuário ainda pode pesquisar através da seção *sinônimos e antônimos* e no dicionário *Aulete Digital*, que traz definições tais como são fornecidas pelos dicionários tradicionais. Além dessas funções, o recurso ainda inclui expressões idiomáticas, figuras, ditos populares, citações e um *link* com os dados trazidos pela enciclopédia *online* Wikipedia.

Em Ruppenhofer et al. (2010), a aproximação da FrameNet com um *thesaurus* é ressaltada quanto a esse objetivo. Assim como nos *thesauri*, a FN

⁸ O Dicionário Criativo é um produto que está sendo desenvolvido pela tese de doutoramento de Felipe Iszlaji de Albuquerque sob orientação do professor Bento Carlos Dias da Silva.

busca reunir aquelas palavras que se agrupam por similaridade semântica. É então de se avaliar que, em alguma medida, os *thesauri* considerem o conceito estruturado pelo *frame*. A principal diferença é que eles não têm como objetivo sistematizar esse construto. Dessa forma, a FrameNet se destaca por ser um amplo dicionário de *frames*. Entretanto, há uma particularidade da FrameNet diante da lexicografia comum: como Fontenelle (2003 apud FILLMORE, 2006, p. 616) assinala, em vez de explorar todos os significados de uma palavra, o interesse é analisá-la em um dos seus sentidos para então definir a qual *frame* ela estaria vinculada.

Nesse sentido, o dicionário eletrônico Copa 2014 se apresenta como produto de inovação tecnológica. Ao utilizar o *frame* como ponto de partida para a organização do léxico, ele oferece um sistema de busca diferente tanto dos dicionários impressos quanto das demais versões eletrônicas, pois permite ao usuário acessar o léxico através de situações, eventos e esquematizações de experiências, representadas pelo *frame*.

2.1.1 Dicionários Impressos vs. Dicionários Eletrônicos

Muitos dicionários impressos, com o passar dos anos, foram se adaptando às novas tecnologias. Tornou-se necessário que versões de papel tivessem uma contraparte eletrônica. O público que começou a se anunciar no século XX e se legitimou no século XXI instaurou uma geração fiel à revolução tecnológica. Lentas transformações foram sendo realizadas. Como aponta Knowles (1990 apud SCHRYVER, 2003, p. 143), o uso de computadores na linguística aconteceu de forma lenta se comparado a outras ciências, como a engenharia, por exemplo. Na elaboração de dicionários, isso também se confirma. Cerquiglini (apud SCHRYVER, 2003, p.143-144) divide essa trajetória nas seguintes etapas: lexicografia assistida por computador; transferência de dicionários de papel para meio eletrônico; e dicionários eletrônicos concebidos a partir de e para tal ambiente.

Com essas três fases, é possível avaliar que a tecnologia garantiu uma nova era para a lexicografia. O uso de computadores foi, aos poucos, adentrando-se no fazer lexicográfico: os dicionários de papel começaram a experimentar as novidades advindas das máquinas, versões eletrônicas foram feitas, e as transformações foram tantas que culminaram em um novo tipo de dicionário. Como ressaltam Usielniewicz

e Olko (2006), o processo contrário, ou seja, criar dicionários impressos a partir dos eletrônicos é um objetivo extremamente difícil, pois, embora tenham sido um aprimoramento das versões impressas, os dicionários eletrônicos mantêm poucos traços em comum com suas origens. Dessa forma, no que tange ao uso de tecnologia na estruturação de dicionários, o processo é irreversível.

As referências cruzadas são um exemplo disso. Como o léxico é rico e mantém diversas relações entre si, não restam dúvidas de que um recurso digital, situado em ambiente de hipertexto, consiga lidar de forma mais eficaz com essa situação. A gama de recursos disponíveis para o meio digital garante grande interatividade com o usuário. Como esses materiais não são limitados pela organização linear, o processo de atualização de dados é garantido de modo mais eficiente que nos recursos impressos, que passam por um longo processo até a comercialização de uma nova edição. Como o perfil do usuário do Copa 2014 são turistas, pessoas envolvidas na organização do evento e a mídia especializada, a busca em um recurso eletrônico *online* responde de forma mais satisfatória do que em um material impresso. Além da oportunidade de atualizar os dados ao longo dos torneios, as pessoas não precisarão levar um dicionário impresso para os jogos ou locais que estiverem visitando, sem contar a chance de os dados serem ofertados de um modo mais dinâmico e que demande menor trabalho do usuário.

Ainda que não restem dúvidas de que o meio digital tenha permitido a operacionalização de várias iniciativas importantes para a estruturação de dicionários, tornando esses recursos atrativos para grande parte da população, isso não implica defender que os dicionários de papel tenham data de validade. Deve-se considerar que esses dicionários, como destacam Schryver (2003), apresentam o valor simbólico de tornar a língua manuseável; consultar livros impressos, muitas vezes, compõe um momento de prazer e descanso, sem contar que não fadiga os olhos como a leitura em monitores. Todavia, há de se avaliar que a revolução tecnológica vivenciada pela lexicografia construiu um novo ambiente de trabalho, que, aos poucos, vai mostrando suas possibilidades, permitindo iniciativas que jamais seriam observadas nos dicionários de papel, um exemplo disso são os vários recursos disponíveis na elaboração de dicionários multilíngues.

2.1.2 Dicionários Eletrônicos Multilíngues

A cada dia, maior é a necessidade de comunicação entre pessoas de diferentes partes do mundo. Seja para passeio, estudo ou negócios, as fronteiras geográficas e/ou culturais não se colocam mais como empecilho para esse contato, o que impõe a necessidade de superar a barreira colocada pelos diferentes sistemas linguísticos. Entretanto, ainda que seja uma tarefa colocada na ordem do dia, diversos desafios são enfrentados quando o objetivo é contrastar línguas com o intuito de sistematizá-las em recursos lexicais. Como lidar com a polissemia em diferentes línguas? O que fazer quando uma palavra não tem uma equivalência estrita nas outras línguas? Como representar um conceito que não é de conhecimento da outra comunidade linguística?

Procurando responder questões dessa natureza, Boas (2009) pontua que a pesquisa em lexicografia computacional, que auxilia na criação de bases de dados lexicais multilíngues, depara-se com desafios mais complexos que aqueles encontrados na estruturação de bases de dados lexicais monolíngues. Por isso, faz-se necessário refletir a respeito de certas questões, como polissemia, padrões sintático-semânticos, padrões de lexicalização e também sobre equivalentes de tradução.

Acerca da polissemia, Boas pondera que seu tratamento em um único idioma já não é uma tarefa simples, e, se o foco passa então a ser a análise da polissemia entre línguas distintas, problemas maiores certamente estarão colocados. Para ele, superar esse problema, além de ser importante para projetos lexicográficos tradicionais, é fundamental para a implementação de bases de dados lexicais multilíngues. Altenberg e Granger (2002 apud BOAS, 2009) expõem que o comportamento da polissemia não é regular, podendo se expressar sob três formas: polissemia sobreposta, polissemia divergente e equivalência inexistente. No primeiro caso, verifica-se a ocorrência de correspondências entre as extensões de significado entre as línguas. Embora esses casos minimizem vários problemas na formalização das entradas lexicais, Boas destaca que esse fenômeno é raro diante das demais ocorrências. Já a segunda situação trata dos dados cujas polissemias apresentam divergências entre as línguas. Um exemplo desse fenômeno é representado pela Unidade Lexical *marcar*. Nos domínios do futebol e do turismo, essa palavra é

polissêmica, associando-se a diferentes situações: por exemplo, *Marcos não marcou o gol; o juiz marcou a falta; Dedé marcou Neymar; marcamos nosso voo com muita antecedência*. Entretanto, no inglês, por exemplo, esses usos terão correspondentes lexicais específicos, como *score*, *assign*, *mark* e *book*, respectivamente, sugerindo que a polissemia ocorra apenas em uma direção. Por último, o terceiro caso expõe situações cujas palavras em uma língua não encontram equivalentes específicos em outra.

Além da sistematização da polissemia, Boas (2009) destaca os padrões de ocorrência verificados em cada língua. Com as informações sobre os diferentes sentidos das palavras, os bancos de dados multilíngues deveriam especificar também seus comportamentos sintáticos. Os padrões de lexicalização são outro componente importante na criação de recursos multilíngues. Conforme Talmy (1985, 2000 apud BOAS, 2009) salienta, as línguas especificam preferências acerca da lexicalização de seus componentes semânticos, isto é, refletem as escolhas dos falantes por determinados usos no processo de atribuição de sentido.

O último desafio explicitado por Boas (2009) na estruturação de bases lexicais multilíngues são as paráfrases e os equivalentes de tradução. O objetivo com essa discussão é mostrar a importância de conhecer em uma língua palavras ou expressões que podem ser substituíveis entre si. Segundo o autor, os padrões sintáticos podem divergir ao evocar um mesmo sentido, por isso, é importante que um recurso lexical multilíngue invista nas paráfrases. Nas palavras de Boas (2009), “quando se trata de equivalentes de tradução, a questão não é somente como medi-los entre as línguas mas também como combiná-los em diferentes paráfrases da língua fonte para diferentes tipos de paráfrases na língua alvo⁹”.

Diante desse rico panorama, a rede semântica FrameNet e a rede lexical WordNet são apresentadas como duas bases de dados multilíngues capazes de viabilizar o alinhamento necessário entre as línguas do dicionário Copa 2014.

⁹ “In other words, when it comes to translation equivalents, the question is not only how to “measure” them cross-linguistically, but also how to match them from different paraphrases in the source language to different types of paraphrases in the target language” (BOAS, 2009, p. 67).

2.2 A FrameNet

Para apresentar a plataforma lexicográfica FrameNet bem como sua expansão para várias línguas, é essencial introduzir a Semântica de *Frames*, teoria que sustenta seus principais postulados. Com tal intuito, essa discussão será abordada em três partes: a inserção da Semântica de *Frames* nos estudos cognitivos da linguagem, o desenvolvimento da FrameNet e, por conseguinte, sua expansão.

2.2.1 A Semântica de Frames nos Estudos Cognitivos da Linguagem

A década de setenta pode ser considerada aquela que iniciou uma nova geração para o cognitivismo linguístico. Diferentes estudos passaram a ser divulgados e debatidos como reação àquilo que até então a academia reconhecia sobre linguagem e cognição. Pesquisadores cujos estudos eram, até então, voltados para o gerativismo, como Ronald Langacker, George Lakoff e Charles Fillmore, passaram a reivindicar a importância da semântica / pragmática para os estudos cognitivos. Com a publicação de trabalhos e a consequente difusão dessas discussões, surgia uma nova linha de pesquisa, designada por seus próprios precursores como Linguística Cognitiva.¹⁰

Antes da sistematização dessas pesquisas, a relação entre linguagem e cognição era analisada de modo idealizado. Na abordagem de Chomsky (1965), defende-se que o ser humano é dotado de um módulo específico para a linguagem, que lhe garante uma pré-disposição para o aprendizado de língua, sendo que o contato com o meio proporciona a atuação desse aparato biológico. Dessa forma, o foco de pesquisa era a *competência*, que discute esse conhecimento inato, pois, de acordo com essa teoria, o *desempenho* dessa capacidade não interfere de modo significativo nesse sistema.

¹⁰ Ainda que pesquisas envolvendo a linguagem e a cognição sejam datadas de antes da década de 1970, o termo Linguística Cognitiva foi utilizado por diferentes estudiosos dessa nova fase de pesquisa e, embora eles não sejam os primeiros a ressaltar a relação linguagem / cognição, o nome vigorou como tal.

Assim sendo, ainda que Chomsky tenha apresentado para a ciência o importante papel da cognição na construção da linguagem, a hipótese da composicionalidade imperou por todos esses anos. Nela, o falante / ouvinte seria guiado pela transparência e pela previsibilidade. A aquisição das regras percebidas no jogo da linguagem dava conta do sistema linguístico; somando-se as partes, chegava-se à compreensão do todo. A metáfora do conduto (REDDY, 1979) representa essa concepção, uma vez que sugere que as palavras atuam como um canal que transporta o sentido. O significado estava, então, contido na forma, e a linguagem era o veículo para tal decodificação. Deste modo, a experiência do indivíduo com o mundo não exercia influência na construção de sentido, já que o processo de significação era externo ao ser humano.

É em contraste a essa teoria que Fillmore (1979) caracterizará o falante / ouvinte gerativista como inocente. O objetivo não só dele como de todos os pesquisadores dessa nova geração é mostrar que o conhecimento das regras não responde a vários fenômenos da língua. Posto que esse falante / ouvinte inocente não consegue fazer qualquer inferência, seria, portanto, incapaz de interpretar metáforas, metonímias e expressões idiomáticas. Todavia, partindo do consenso de que esses processos, salvo exceções, são prontamente compreendidos, era necessário algo mais que a soma das partes. Exatamente por isso, explorar os fenômenos a que a ciência, até aquele momento, não havia se dedicado passa a ser o cerne para os estudos cognitivistas da linguagem. O foco, então, passa a se direcionar para o falante / ouvinte real.

Em *Philosophy in The Flesh*, Lakoff e Johnson (1999) exploram as premissas que dão vida à ciência cognitiva: “a mente é inerentemente corpórea”; “o pensamento é amplamente inconsciente”; “conceitos abstratos são largamente metafóricos”. Com esses pontos de partida, pretende-se explicar a influência da experiência sensório-motora e das estruturas neurais nos processos de conceptualização e categorização envolvidos na linguagem. Para esse empreendimento, o enfoque é amplo, o que torna a Linguística Cognitiva um ramo de pesquisa heterogêneo, que preserva relação com outras áreas, como a Biologia e a Inteligência Artificial, e se organiza em várias vertentes para investigar a linguagem, como a Semântica de *Frames* (FILLMORE, 1982), a Teoria da Metáfora Conceptual (LAKOFF & JONHSON, 1980) e a Gramática das Construções (FILLMORE, 2008a; GOLDBERG, 1995, 2006).

O desafio aceito é, sem dúvida, ousado, uma vez que investigar certos fenômenos envolvidos na linguagem humana passa por rever a filosofia vigente por todos esses anos. Como Lakoff (1999) ressalta, essa iniciativa pode fazer com que a filosofia jamais seja a mesma. Para Salomão (2010), a “virada cognitiva”, datada no fim da década de cinquenta do século passado, quando o computador colocava em xeque o que se entendia por conhecimento, já estava causando uma irreversível mudança de paradigma na forma como a linguagem até então era entendida.

É nesse contexto que surge a Semântica de *Frames* (FILLMORE, 1982, 1985; PETRUCK, 1996), uma abordagem que enfatiza a estreita relação linguagem / experiência. De acordo com seu precursor, Charles J. Fillmore, o objetivo é buscar as razões que impulsionam a categorização humana. Para tal, a descrição da língua se dá a partir do uso, sendo considerado um programa de semântica empírica e um modelo descritivo para apresentar os resultados das análises. Nas palavras de Fillmore,

com o termo *semântica de frames*, eu tenho em mente um programa de pesquisa em linguística empírica e uma metodologia descritiva para apresentar os resultados de tal pesquisa. (...) Pelo termo *frame*, eu tenho em mente qualquer sistema de conceitos relacionado de tal forma que, para entender qualquer um deles, você tem de entender toda a estrutura na qual ele se encaixa; quando um dos conceitos em dada estrutura é introduzido dentro de um texto ou de uma conversa, todos os outros são automaticamente disponibilizados. (FILLMORE, 1982, p.111).¹¹

Quando optou pela palavra *frame*, Fillmore trouxe para a linguística um termo utilizado na Inteligência Artificial (MINSKY, 1987) e também na sociologia (GOFFMAN, 1974). Para Minsky, esse termo mostra que existem estruturas de dados estereotipadas que representam as situações, sendo o exemplo da festa de aniversário utilizado pelo autor para ilustrar essa discussão. Ele pondera que as definições de dicionário nunca dizem o suficiente. Ainda que qualquer pessoa saiba que uma festa desse tipo envolve mais que um encontro para comemorar o aniversário de alguém, nenhuma definição breve consegue atingir a complexidade de tal evento. Ou seja, se algum dos organizadores lamenta que se esqueceu da vela, dificilmente, alguém irá questionar se a luz acabou. Já Goffman, por um viés

¹¹ "With the term 'frame semantics' I have in mind a research program in empirical semantics and a descriptive framework for presenting the results of such research...By the term 'frame' I have in mind any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available".

sociológico, usa a palavra *frame* para enfatizar as intenções, as perspectivas, os rituais e os padrões que os indivíduos estabelecem na interação cotidiana. De acordo com ele, a atuação do Homem na sociedade pode ser comparada a uma peça de teatro: nela, assumem-se vários papéis a depender da função da interação estabelecida.

Em diálogo com esses autores, Fillmore irá defender que o processo de significação das palavras está diretamente ligado à compreensão das instituições sociais ou da estrutura de experiência que elas pressupõem. Deste modo, assumindo que as palavras estejam ligadas às experiências humanas, ele assinala que os *frames* consistem, em muitos casos, em uma porção significativa de cultura (FILLMORE, 1982, p. 118).

Antes de chegar a essa definição, que associa o conhecimento linguístico ao conhecimento enciclopédico, Fillmore utilizou esse termo considerando enfoques específicos (FILLMORE, 1982, 2006, 2008). A distinção entre *frame* cognitivo e *frame* linguístico, por exemplo, denotava a tentativa de separar o conhecimento necessário para entender as estruturas de conceito e as formas linguísticas que evocam tais estruturas. Semelhante foi a separação entre *cena* e *frame*,¹² já que dividia as estruturas de conhecimento, crenças e experiências do ser humano das escolhas linguísticas para compor determinada situação, como palavras e categorias gramaticais. Como essas estruturas são compreendidas como construtos mentais, não são, necessariamente, manifestadas linguisticamente. Faz-se, então, a distinção entre *frames* invocados e evocados:

Frames interpretativos podem ser introduzidos no processo de compreensão de um texto sendo invocados pelo intérprete ou sendo evocados pelo texto. Um *frame* é invocado quando o intérprete, no processo de construção de sentido de dado segmento textual, é capaz de fazer uma interpretação situando o conteúdo textual em um padrão que é conhecido independentemente do texto. Um *frame* é evocado pelo texto se alguma forma ou padrão linguístico é convencionalmente associado com o *frame* em questão. (FILLMORE, 1985, p.232)¹³

Aqueles que são invocados contam com pistas textuais, sejam escritas ou faladas, para depreenderem o conteúdo da informação, porém não estão ancorados

¹² A distinção *cena* e *frame* é pormenorizada ao decorrer deste trabalho devido ao fato de Schmidt, na estruturação do dicionário *Kicktionary*, retomar essa terminologia e, conseqüentemente, a teoria ao qual está vinculada.

¹³ "Interpretative frames can be introduced into the process of understanding a text through being invoked by the interpreter or through being evoked by the text. A frame is invoked when the interpreter, in trying to make sense of a text segment, is able to assign it an interpretation by situating its content in a pattern that is known independently of a text. A frame is evoked by the text if some linguistic form or pattern is conventionally associated with the frame in question".

em nenhum item lexical específico. Por exemplo, considerando a sentença (1), não se encontram elementos lexicais indicadores de que o interlocutor esteja em uma atividade turística, entretanto algumas informações sugerem que se trata de tal evento. Quando o *frame* é evocado, há palavras que o explicitam. Na sentença (2), o texto é construído a partir dessas palavras, que evidenciam o evento turístico.

- (1) Não há nada mais prazeroso que chegar a um lugar que ninguém te conhece. Você acorda bem cedo, vai à praia e assiste ao pôr do sol, deixando todas as obrigações de lado por uns dias.
- (2) Sempre que chegar a uma cidade turística, tome cuidado com a bolsa e demais pertences. Os assaltantes reconhecem os visitantes de longe.

De modo geral, o turista é aquele que vai a lugares cujos habitantes e ambiente não conhece. Além disso, o período escolhido para fazer turismo, geralmente, é aquele em que o indivíduo pode se ausentar das tarefas profissionais. Dessa forma, em (1), *frames* do domínio do turismo são invocados pelo texto, possibilitando que o leitor compreenda o objetivo comunicativo do texto. Já, na sentença (2), o texto se estrutura através da compreensão de expressões específicas do domínio turístico, como *cidade turística* e *visitante*.

A relação da semântica com a pragmática, nesse sentido, é destacada. Quando questionado sobre quais motivos justificam a escolha do nome de sua teoria – Semântica de *Frames*, e não Pragmática de *Frames* –, Fillmore (FILLMORE & ANDOR, 2010) ressalta que a semântica trata da forma linguística e das convenções estabelecidas pela comunidade que compartilha interpretações dos usos de determinadas formas, já a pragmática trata das configurações de interpretação, sinalizando como são as convenções que levam a ela. Segundo ele, não há semântica que seja completamente livre da pragmática, no entanto há, sim, pragmática livre de referências específicas a formas linguísticas. Nesse sentido, enquanto o exemplo (2) explicita formas linguísticas específicas na estruturação de *frames*, o exemplo (1) os invoca a partir da compreensão da mensagem veiculada.

Com esse aparato de postulações teóricas, a literatura da Semântica de *Frames* encorajou Fillmore a desenvolver a FrameNet, que pode ser considerada uma extensão de sua teoria para o domínio da lexicografia computacional. A seção seguinte tratará desse empreendimento.

2.2.2 Desenvolvimento da Rede Semântica

A FrameNet é um projeto que surgiu em 1997, no International Computer Science Institute (ICSI), sob liderança de Fillmore, em Berkeley, Califórnia. O interesse inicial era criar um recurso lexical que descrevesse a língua inglesa por um viés semântico e sintático. A metodologia seria guiada pela teoria da Semântica de *Frames* (FILLMORE, 1982, 1985), e as análises, sustentadas por evidências em *corpora*. Com o desenvolver dos resultados, essa ferramenta passou também a ser útil para profissionais ligados à lexicografia, como alternativa na estruturação de dicionários, e para a Linguística Computacional, no desenvolvimento de tarefas ligadas ao Processamento de Linguagem Natural (PLN).

Conforme avalia Atkins em trabalhos colaborativos (ATKINS ET AL., 2003; ATKINS & RUNDEL, 2008), a base de dados da FN apresenta potencial para apoiar, acelerar e enriquecer a lexicografia tradicional. Como as escolhas acerca das informações que estarão nas entradas lexicais de um dicionário requerem um olhar atento do lexicógrafo, os autores afirmam que, de modo imediato, o uso de *corpus* deva ser a contribuição mais relevante desse projeto para a lexicografia (ATKINS & RUNDEL, 2008, p. 147).

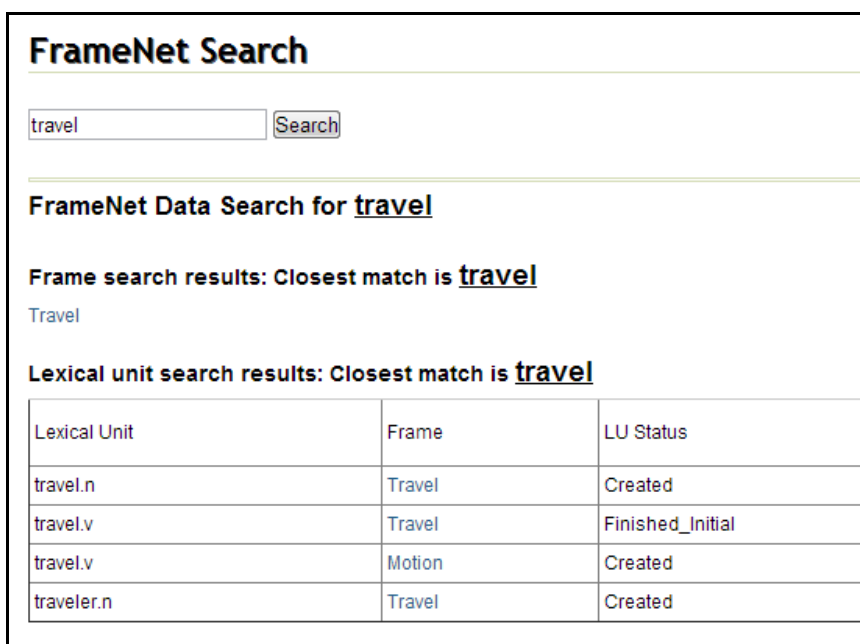
Toda análise lexical realizada na FrameNet é apoiada em textos autênticos, advindos do British National Corpus (BNC), American National Corpus (ANC) e Wall Street Journal (WSJ). São mais de 100 milhões de palavras que fornecem subsídio para as análises. A opção por utilizar as ferramentas da Linguística de *Corpus*, submetendo cada generalização a dados reais, faz com que a FrameNet reitere a proposição da Semântica de *Frames* de fornecer um programa de pesquisa em semântica empírica para a descrição da língua. Por esse motivo, Atkins e Rundel (2008) ressaltam que essa teoria oferece uma possibilidade sistemática de garantir que as informações relevantes sejam capturadas.

O banco de dados da FrameNet encontra-se disponível *online* no sítio <<https://framenet.icsi.berkeley.edu/>>. ¹⁴ Grosso modo, tal banco é composto por três grandes tipos de dados: *Frames*, Unidades Lexicais e Sentenças. No que dizem respeito aos *frames*, eles podem ser entendidos como a modelagem computacional de uma estrutura de conhecimento reconhecível em uma dada cultura. Contam com

¹⁴ O banco de dados da FrameNet está em constante atualização. Última visualização 10/01/13.

uma definição e com a especificação dos atores, ferramentas e circunstâncias que os constituem (os Elementos de *Frame*). Já as Unidades Lexicais (ULs) são entendidas como o pareamento de uma forma, com todas as suas flexões, a um significado específico, ou seja, é um lexema pareado a um *frame*. Seu papel é importante porque, a partir da sua inserção, serão delimitados quais serão os Elementos de *Frame* em sua estrutura de valência. Além dos verbos, classe de palavra considerada prototípica no processo de estruturação de *frames*, a FrameNet também analisa nomes, adjetivos, advérbios e preposições atuando como alvo. Por fim, as sentenças, quando anotadas, fornecem evidência empírica para as análises que levaram à constituição dos *frames* e definição das ULs. Atualmente, há mais de 12.000 Unidades Lexicais no banco de dados da FrameNet, das quais mais de 8.000 estão totalmente analisadas obedecendo aos critérios adotados pelo projeto. Esses dados se distribuem em mais de 1.100 *frames*, que são exemplificados nas 170 mil sentenças anotadas.

Assim como nos outros dicionários eletrônicos, a FrameNet apresenta uma ferramenta, **FrameNet Search**, que possibilita que o usuário digite sua busca. O resultado é apresentado em termos de *frames* e Unidades Lexicais. A busca por *travel*, Figura 9, retornou quatro Unidades Lexicais: dois lexemas nominais *travel* (*viagem*) e *traveler* (*viajante*) e um verbal *travel* (*viajar*), associados ao *frame* Travel (*Viajar*); e um lexema verbal *travel* (*viajar*), vinculado ao *frame* Motion (*Movimento*).



The screenshot shows the 'FrameNet Search' interface. At the top, there is a search bar with the word 'travel' entered and a 'Search' button. Below the search bar, the results are displayed under the heading 'FrameNet Data Search for travel'. The first section is 'Frame search results: Closest match is travel', which lists 'Travel'. The second section is 'Lexical unit search results: Closest match is travel', which contains a table with the following data:

Lexical Unit	Frame	LU Status
travel.n	Travel	Created
travel.v	Travel	Finished_Initial
travel.v	Motion	Created
traveler.n	Travel	Created

Figura 9: Resultado da busca por *travel* na ferramenta *FrameNet Search*
 Fonte: <https://framenet.icsi.berkeley.edu/fndrupal/framenet_search>

Outra opção de busca é o **Frame Index**, Figura 10. Esse recurso permite que o usuário tenha acesso a todos os *frames* da plataforma. A busca por `Travel` trouxe sua definição bem como Elementos de *Frame*, Unidades Lexicais e relações entre *frames*. A disposição desses *frames* nesse recurso se dá por ordem alfabética, entretanto a prioridade é, quando possível, uni-los por similaridade semântica. Por exemplo, os *frames* relacionados a atividades como `Activity_finish`, `Activity_ongoing` e os demais mostram esse intuito. Como apresentam similaridade semântica e a palavra *activity* em comum, optou-se por agrupá-los.

The screenshot shows the 'Frame Index' search results for the term 'Travel'. On the left, there is a list of frames under the heading 'Frame Index', with a navigation bar for letters A through Z. The main content area is titled 'Travel' and includes a 'Definition:' section with a detailed explanation of the frame. Below the definition are three example sentences with highlighted frame elements: 'Ellen JOURNEYED to Europe with five suitcases.', 'Samantha JOURNEYED 2500 miles with her family by sea to China.', and 'The Osbournes took a TRIP from Beverly Hills to London on the Concorde.'. There are also sections for 'FEs:', 'Core:', and 'Area [Area]' with a description of the area frame element.

Figura 10: Resultado da busca por `Travel` na ferramenta *Frame Index*

Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=frameIndex>>

O usuário que tiver interesse sobre a entrada lexical pode se dirigir diretamente ao **Lexical Unit Index**, Figura 11. Essa ferramenta, além de trazer o link para o *frame* de cada UL, apresenta outras duas opções de busca: **Lexical Entry**, Figura 12, e **Annotation**, Figura 13. A finalidade é que o usuário possa conferir como se deu o comportamento semântico e sintático de cada Elemento de *Frame*.

FrameNet Index of Lexical Units

This page is an index to alphabetical lists of the names of the lexical units (LUs).

Each LU name is followed by the part of speech, the name of the relevant frame, and its status. If a lexical unit has annotated in FN2) or "FN1_sent" (meaning annotated in FN1), it will be followed by links to the HTML files for the units on which work has not been completed may have only a link for the lexical entry, or no link at all. The lexical LU:Frame Elements and their Syntactic Realizations; and Valence Patterns.

travel

<#> | [A](#) | [B](#) | [C](#) | [D](#) | [E](#) | [F](#) | [G](#) | [H](#) | [I](#) | [J](#) | [K](#) | [L](#) | [M](#) | [N](#) | [O](#) | [P](#) | [Q](#) | [R](#) | [S](#) | [T](#) | [U](#) | [V](#) | [W](#) | [X](#) | [Y](#) | [Z](#) | [All](#)

T

- TV room.n ([Building subparts](#)) [Created Lexical Entry](#)
- Thursday.n ([Calendric unit](#)) [BTDT Lexical Entry Annotation](#)
- Tuesday.n ([Calendric unit](#)) [BTDT Lexical Entry Annotation](#)
- Turk.n ([People by origin](#)) [Created Lexical Entry Annotation](#)
- t-shirt.n ([Clothing](#)) [Finished Initial Lexical Entry Annotation](#)
- tabernacle.n ([Buildings](#)) [Finished X-Gov Lexical Entry Annotation](#)
- tablespoon.n ([Measure volume](#)) [Finished Initial Lexical Entry Annotation](#)
- taciturn.a ([Volubility](#)) [Finished Initial Lexical Entry Annotation](#)
- tack.n ([Connectors](#)) [Finished Initial Lexical Entry Annotation](#)
- tack.v ([Operate vehicle](#)) [FN1_Sent Lexical Entry Annotation](#)
- tack.v ([Attaching](#)) [Finished Initial Lexical Entry Annotation](#)
- tack.v ([Self motion](#)) [Created Lexical Entry](#)
- tack_on.v ([Attaching](#)) [Created Lexical Entry](#)
- tactful.a ([Social interaction evaluation](#)) [Finished Initial Lexical Entry Annotation](#)
- tactic.n ([Means](#)) [Created Lexical Entry Annotation](#)
- tag along.v ([Cotheme](#)) [Created Lexical Entry](#)
- taiga.n ([Biological area](#)) [Finished Initial Lexical Entry Annotation](#)

Figura 11: Resultado da busca por *travel* na ferramenta *Lexical Unit Index*
Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=lulnde>>.

Em *Lexical Entry*, é sugerida uma definição para a Unidade Lexical e são exibidos quais foram os padrões de valência semânticos e sintáticos que os Elementos de *Frame* seguiram nos *corpora*.

Frame Element	Number Annotated	Realization(s)
Area	(25)	AVP.Dep (7) DNI.-- (1) INL.-- (3) NP.Obj (1) PP[around].Dep (6) PP[in].Dep (3) PP[inside].Dep (1) PP[round].Dep (1) PP[through].Dep (1) PP[within].Dep (1)
Baggage	(2)	PP[with].Dep (2)
Co-participant	(9)	PP[with].Dep (8) AVP.Dep (1)
Distance	(3)	AVP.Dep (1) PP[than].Dep (1) NP.Dep (1)
Duration	(4)	PP[for].Dep (3) PP[over].Dep (1)

Figura 12: Resultado da busca por *travel* na opção *Lexical Entry*
Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=lulIndex>>.

O usuário pode ainda optar por visualizar como os EFs aparecem nas sentenças analisadas – Figura 13. Em caixa alta, é destacado o alvo, como também é conhecida a Unidade Lexical; os sintagmas evidenciados entre colchetes são os EFs, apresentados com seus respectivos nomes.

- 429-s20-rcoll-aircraft
 1. [TravelerHer Royal Highness] TRAVELLED^{Target} [Mode_of_transportationin an aircraft of The Queen 's Flight] .[GoalINI]
- 429-s20-rcoll-country
 1. If [Traveleryou] TRAVEL^{Target} [Goalto a tropical country] , write to that country 's tourist office to tell them that you are concerned about
 2. [TravelerPaul Pratt] TRAVELLED^{Target} [Paththrough forty-eight countries] .
 3. [TravelerHe] has TRAVELLED^{Target} [Co_participantwith the team] [Goalto every country in Europe] and his outstanding memory was Ireland
 4. [TravelerCustoms officers from each country] would be allowed to TRAVEL^{Target} [Goalto either country] [Purposein pursuit of suspected t
 5. [TravelerThe Essex Section] will TRAVEL^{Target} [Mode_of_transportationby coach] [Goalto the poppy country of Picardy , Northern France] in Amiens and district] .
- 429-s20-rcoll-world
 1. Although [Travelerhe] has TRAVELLED^{Target} [Paththe world] , Edinburgh and the family memories it evokes , are still profoundly importan
 2. [TravelerThe owner , who had recently died] , had TRAVELLED^{Target} [Paththe world] .
 3. Carefully planning the cheapest air fares , allocating her meagre budget as sparingly as possible on food and accommodation , [Travelershe]
 4. [TravelerShe] TRAVELLED^{Target} [Pathfound the world] , paid two visits to the United States , and for a time taught English in Japan .
 5. He 'll read extracts from the works of [TravelerAmerican writers] [Travelerwho] TRAVELLED^{Target} [Goalto Paris] [Timeafter the first world
- 680-s20-pass
 1. Days began early and ended late so that [Pathmaximum distances] could be TRAVELLED^{Target} .[TravelerCNI]
- 710-s20-adverbabroad
 1. After a five-year apprenticeship to Richard Lane of Manchester , [Travelerhe] TRAVELLED^{Target} [Areaabroad] [Period_of_itationsfor a year architectural details] .
 2. [Period_of_itationsFrom 1802 to 1810] [Travelerhe] TRAVELLED^{Target} [Placeabroad] , [Goalto Berlin , Dresden , Naples , Rome , Milan , a
 3. [TravelerThe two girls who shared the upstairs flat] were models , and [Frequency-often] TRAVELLED^{Target} [Areaabroad] [Purposeon assignm
 4. It meant that [Travelershe] could see you every day , work closely with you , even TRAVEL^{Target} [Areaabroad] [Co_participantwith you] . '
 5. [TravelerThey] can TRAVEL^{Target} [Areaabroad] going to shows and fights .
- 710-s20-adverbtogether
 1. [TravelerWe] could TRAVEL^{Target} together [Goalas far as Doha] and I would go on alone to Abu Dhabi .
 2. [TravelerWe] were going to TRAVEL^{Target} together [Goalas far as Paris] , I was to go on to Calais .

Figura 13: Resultado da busca por *travel* na opção *Annotation*

Fonte: < <https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=luIndex> >

Outra ferramenta disponibilizada é o **FrameGrapher**. Ela sinaliza quais relações são notadas entre os *frames* e viabiliza as relações entre os elementos do *frame*. Cada cor de seta representa uma relação, por exemplo, na Figura 14, *Travel* é ligado a *Self_motion* (Auto_movimento) através da seta vermelha, que indica relação de herança; a seta azul, que liga *Travel* a *Setting_out* (Partir), exhibe a relação de subframe.

Para conhecer as relações entre os EFs, é necessário clicar na seta de cada relação desejada. O objetivo desse instrumento é exibir como o parentesco entre os *frames* reflete as relações entre seus elementos. Na Figura 15, evidenciam-se as relações entre os EFs de *Self_motion* e *Travel*. As indicações *c* e *nc* referem-se aos elementos considerados *core* (nuclear) e *non core* (não nuclear), respectivamente. Como a relação entre esses *frames* é de herança, os elementos nucleares, conforme previsto pela FN, são preservados de alguma forma do *frame* mãe para o filho.

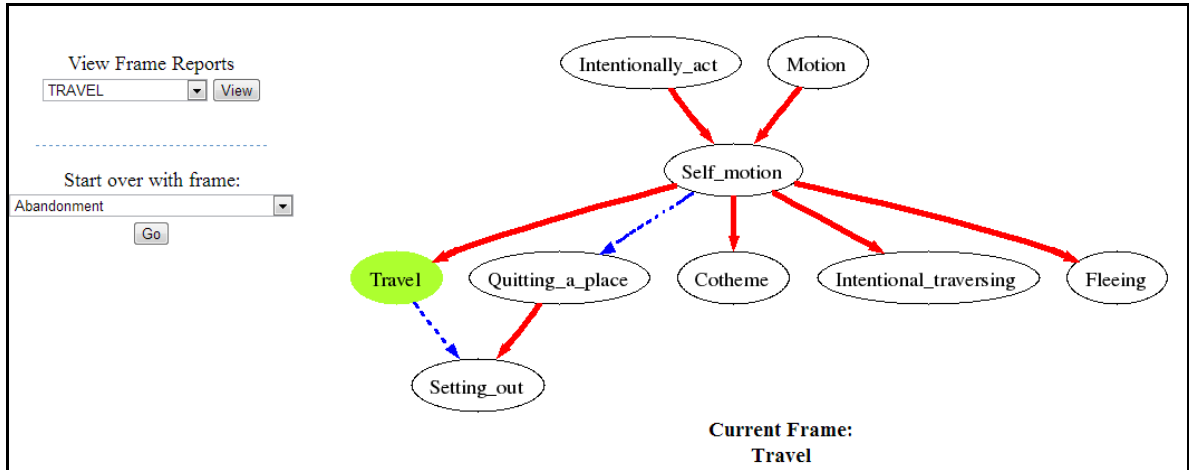


Figura 14: Resultado da busca por *Travel* na ferramenta *FrameGrapher*
 Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/FrameGrapher>>

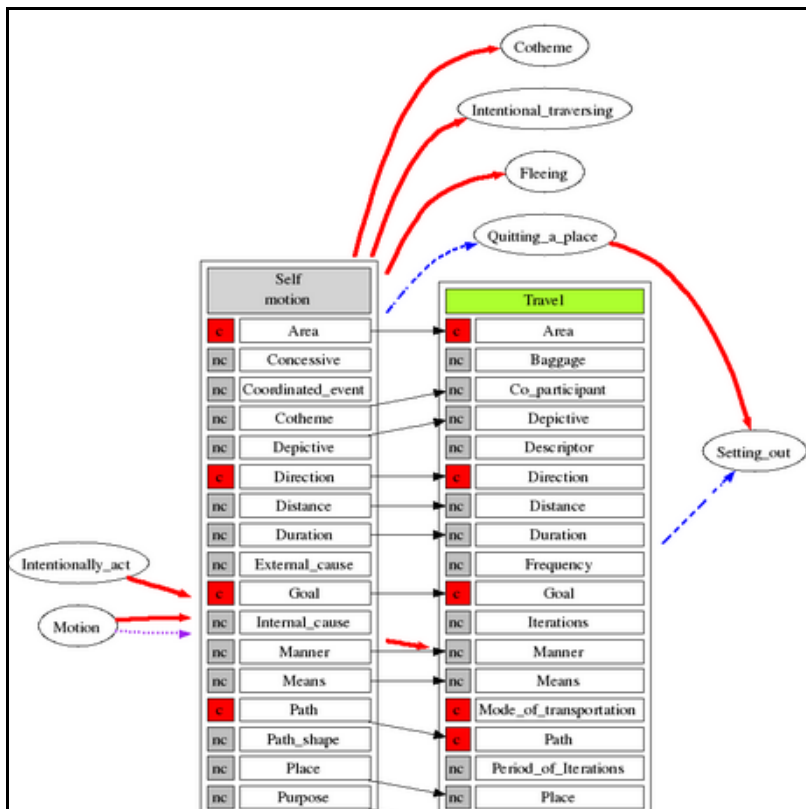


Figura 15: Relação entre os Elementos do *Frame* *Self_motion* e *Travel*
 Fonte: idem

Embora haja traços de uma organização linear, que prioriza a ordenação alfabética, a estruturação da *FrameNet* permite alterações e inclusões em seu banco de dados. Boas (2009) destaca que uma vantagem de esse dicionário não ser completamente linear é a possibilidade de reavaliar as análises de acordo com os resultados sugeridos pelos *corpora*. Visando a esse objetivo, Petruck et al. (2004) pontuam a importância do *software* que permite esse tipo de tratamento. Quando os

frames são reavaliados, um olhar mais detalhado garante uma análise mais refinada, que gera novos *frames*, redireciona ULs etc. A realização desse procedimento é chamada pelos autores de *reframing*.

A metodologia utilizada para analisar os dados é apresentada pelo *The Book* (RUPPENHOFER et al., 2010), obra organizada pela equipe da FrameNet. O ponto de partida é o *frame*. Desse modo, o objetivo da FrameNet é mapear o comportamento dos *frames*, investigando as Unidades Lexicais que os suscitam e os elementos que os compõem. Para estruturar um *frame*, é preciso conceber como a cena se constitui. Na Figura 16, é apresentada a definição de *Travel*, que prevê um viajante indo para alguma atividade, geralmente planejada com antecedência, na qual ele se move de um local de origem para o local desejado. Há algum meio de transporte que direciona o viajante para o local almejado; a viagem pode ser feita com acompanhantes e bagagens; geralmente, há uma duração pré-determinada, assim como a estadia. Sentenças, como [*Ellen* *VIAJANTE*] *VIAJOU* ^{ALVO} [*para a Europa* *DESTINO*] [*com cinco malas* *BAGAGEM*], adaptada do inglês, Figura 16, são exibidas com o intuito de ilustrar o *frame*.

Travel

Definition:

In this frame a **Traveler** goes on a journey, an activity, generally planned in advance, in which the **Traveler** moves from a **Source** location to a **Goal** along a **Path** or within an **Area**. The journey can be accompanied by **Co participants** and **Baggage**. The **Duration** or **Distance** of the journey, both generally long, may also be described as may be the **Mode of transportation**. Words in this frame emphasize the whole process of getting from one place to another, rather than profiling merely the beginning or the end of the journey.

Ellen **JOURNEYED** **to Europe** **with five suitcases**.

Samantha **JOURNEYED** **2500 miles** **with her family** **by sea** **to China**.

The Osbournes **took a TRIP** **from Beverly Hills** **to London** **on the Concorde**.

Figura 16: Definição do *frame* *Travel*

Fonte: <<https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Travel>>

A definição de um *frame* é resultado de um estudo que avalia quais propriedades são necessárias para esquematizar uma dada situação. As marcações coloridas destacam os elementos que compõem os *frames*. Como Salomão (2009) explica, os Elementos de *Frame*, de acordo com uma designação mais tradicional da literatura, são remetidos às Funções Temáticas da grade argumental dos núcleos lexicais, nomeadas também como Papéis Temáticos. Se o intuito for traçar uma “perspectiva genealógica”, esses elementos são herdados dos Casos Semânticos, propostos por Fillmore (1968) na Gramática de Casos. Atualmente, a FrameNet trata os EFs como funções microtemáticas (FILLMORE, 2003 ; SALOMÃO, 2009). O

propósito é garantir que as análises serão mais fiéis às funções semânticas dos argumentos no discurso. O Agente, por exemplo, embora consiga se distinguir de um Experienciador, é um papel temático amplo, que agrega argumentos com diferentes atuações.

- (3) Pedro visitou seu pai quando ele estava doente na UTI
- (4) Maria visitou o Pão de Açúcar quando esteve no Rio de Janeiro.

Em (3) e (4), os sintagmas nominais *Pedro* e *Maria* contemplam as propriedades necessárias para serem admitidos como Agente, já que ambos são seres animados que efetuam alguma ação. Entretanto, os contextos em que se inserem requerem características particulares. Quando a FrameNet escolhe nomes mais específicos para os EFs, o objetivo é registrar que as propriedades de cada constituinte interferem na construção do *frame*. Assim, *Visiting* (Visitar) denomina o EF com função agentiva como Agente, já o *Touring*, *frame* que trata de um tipo específico de visita, identifica o elemento com função agentiva como Turista. Dessa forma, os sintagmas *Pedro* e *Maria* são identificados como Agente e Turista, respectivamente.

No processo de caracterização de um *frame*, é também de se destacar a importância dos argumentos internos. Nos exemplos (3) e (4), o valor assumido pelos sintagmas nominais *seu pai* e *Pão de Açúcar* é de fundamental importância; nesse caso, eles precisam em qual *frame* cada um estará inserido. Na FN, os nomes deles são, respectivamente, Entidade e Atração. Embora Atkins e Rundel (2008) tenham ressaltado que as informações acerca dos complementos sejam, para as entradas de dicionário, mais importantes que o sujeito, casos como (5) e (6) destacam a relevância desse sintagma no processo de definição de *frame*, e, conseqüentemente, nas entradas do dicionário Copa 2014.

- (5) Dunga substituiu o atacante Robinho no último jogo da seleção.
- (6) Nilmar substituiu o atacante Robinho no último jogo da seleção.

Todos aqueles que tiverem algum conhecimento dos *frames* que constroem o domínio do futebol saberão que as duas sentenças não são similares. Em (5), o sintagma nominal *Dunga* não atua da mesma forma que *Nilmar*, em (6). Isso ocorre porque *Dunga* representa o técnico do time e, por isso, pode nomear substituições. Já *Nilmar* é um jogador, indicando que ele ocupou o lugar do atacante Robinho.

Assim sendo, Salomão (2009), a partir desses exemplos, pondera que situações dessa natureza ainda não são satisfatoriamente respondidas. Todavia, se a estruturação do *frame* que trata da partida de futebol visar a aprofundar cada componente desse cenário, certamente, haverá um *frame* de substituição composto por duas perspectivas: a do técnico e a do jogador, respondendo assim à reflexão exposta em Salomão (2009).

FEs:	
Core:	
Area [Area] Semantic Type: Location	This is the Area in which the traveling takes place. This frame element describes the enclosed area inside which travelling, of unspecified Source , Path or Goal takes place. We TRAVELLED in Europe.
Direction [dir] Excludes: Area	The direction in which the Traveler goes. They began their ODYSSEY north.
Goal [Goal] Semantic Type: Goal Excludes: Area	The Goal is the location where the travelers end up.
Mode of transportation [MoT]	The Mode of transportation expresses how the motion of the Traveler is effected, by their body or by a vehicle which holds and conveys the Traveler . Vehicles can move in any way and in any medium. They are usually expressed obliquely with 'in' or 'by'. Barney used to TRAVEL by bus a lot Strom TRAVELLED on foot to see the Pope
Path [Path] Semantic Type: Path Excludes: Area	The Path is the route along which the travel takes place.
Source [Src] Semantic Type: Source Excludes: Area	The Source is the starting point of the trip.
Traveler [Trav] Semantic Type: Sentient	This is the living being which travels. Normally, the Traveler is expressed as an external argument.

Figura 17: Definição dos Elementos de *Frame* nucleares em *Travel*

Fonte: <<https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Travel>>

De acordo com os procedimentos da FrameNet, os Elementos de *Frame* podem ser nucleares, periféricos e extratemáticos. Os nucleares são elementos conceptualmente necessários para a constituição de um evento, e, a partir dele, um *frame* se distingue dos demais. Na Figura 17, há a definição dos elementos nucleares de *Travel*, que são: *Area* (Área), *Direction* (Direção), *Goal* (Destino), *Mode_of_transportation* (Meio_de_transporte), *Path* (Trajetória), *Source* (Origem) e *Traveler* (Viajante). A designação deles como EFs nucleares significa que o evento da viagem inclui, necessariamente, esses participantes. Ou seja, a concepção desse evento requer dados como viajante, lugar de origem e lugar de chegada, por exemplo.

Os Elementos de *Frame* não centrais, conforme a FN prevê, são divididos entre periféricos e extratemáticos – Figura 18. Os periféricos atuam para acrescentar informações que se vinculam à estrutura do *frame*. Entretanto, são dispensáveis

para a sua constituição. Em relação ao *frame* *Travel*, elementos como os que especificam bagagem e duração da viagem não determinam a composição desse *frame*, mas identificam informações auxiliares à ideia principal, atuando, por isso, como EFs periféricos. Há casos em que elementos periféricos de um *frame* são também encontrados em outros, como os EFs Tempo, Lugar e Duração, que participam de vários *frames*.

Non-Core:	
Baggage [Bag]	The Baggage are the items necessary for travel that accompany the Traveler . Ellen JOURNEYED to Europe with five suitcases .
Co-participant [cop] Semantic Type: Sentient	The Co-participant is the person or persons who accompany the Traveler on the journey.
Depictive [Depict] Semantic Type: State	The state of the Traveler during the journey. We TRAVELED around unencumbered .
Descriptor []	A characteristic of the traveling event.
Distance [Dist] Semantic Type: Quantity	This FE identifies the Distance traveled.
Duration [Dur] Semantic Type: Duration	This FE identifies the Duration of time during which the trip occurs.
Frequency [Freq]	The Frequency with which the Traveler makes the journey.
Iterations [Iter]	The number of times the trip is traveled by the Travelers .
Manner [Manr] Semantic Type: Manner	The Manner in which the traveling takes place. The Brownie troop TRAVELLED hurriedly .

Figura 18: Definição dos Elementos de *Frame* não nucleares em *Travel*

Fonte: <<https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Travel>>

Já os Elementos de *Frame* extratemáticos se combinam com vários *frames*, incorporando contextos maiores de atuação. A diferença deles para os periféricos é que realçam aspectos fora do escopo constituído pelo *frame*; isso acontece porque inserem atributos próprios de outros *frames*, nos quais podem chegar a atuar como unidades alvo. Na sentença, [*Clara VIAJANTE*] *VIAJA*^{ALVO} [*frequentemente FREQUÊNCIA*] [*para Florianópolis DESTINO*], o EF Frequência atua como extratemático, sinalizando que, embora a estrutura de *Travel* o aceite como EF, ele evoca um *frame* próprio, o *Frequency* (Frequência), no qual há diversas Unidades Lexicais que designam essa função, dentre elas o advérbio *frequently*.

Esse comportamento irregular, ressaltado no apêndice A do *Book* (cf. RUPPENHOFER ET AL., 2010), é evidenciado através dos exemplos (7), (8) e (9) extraídos do banco de dados da FrameNet e traduzidos para o português.

- (7) [Dois anos mais tarde TEMPO], [o petróleo NOVO] SUBSTITUIU^{ALVO} [o carvão VELHO] [como opção de energia PAPEL].

- (8) [Um estado australiano _{AGENTE}] propôs ao país adotar uma declaração de direitos e SUBSTITUIR^{ALVO} [a rainha _{VELHO}] [como chefe de estado _{PAPEL}] [por um presidente australiano _{NOVO}].
- (9) [Este livro em particular _{ENTIDADE_CRIADA}] foi CRIADO^{ALVO} [como um texto introdutório _{PAPEL}].

As três sentenças possuem o EF *Role* (Papel), porém, em cada uma, esse EF assume uma função específica. Em (7), o *frame* evocado é *Take_place_of* (Ocupar_lugar), nele, o EF Papel é considerado nuclear; em (8), o *frame* em questão é *Replacing* (Substituição), e o EF Papel atua como periférico; por último, em (9), esse EF participa do *frame* *Intentionally_create* (Criar_intencionalmente) como EF extratemático. Situações como essas demonstram que um EF considerado extratemático pode assumir valores distintos a depender do *frame* considerado. Daí, um EF extratemático em um *frame* admitir valor de EF nuclear ou periférico em outros. As mesmas considerações são feitas para outros EFs como Razão e Descrição, que também atuam fora da grade argumental em diversos *frames*. Além desses, há ainda um tipo de Elemento de *Frame*, chamado de *core unexpressed* (nuclear não expresso), que marca uma exceção na relação Herança, na medida em que, apesar de serem anotados no *frame* mãe, tais EFs podem não ser expressos nos herdeiros.

Fenômenos também previstos pela FrameNet ocorrem quando os Elementos de *Frame* atuam de maneira interdependente: as relações *coreset*, *exclui* e *requer* ilustram essa característica. A relação *coreset* ou *coreness* acontece quando um *frame* apresenta EFs que agem em conjunto, de modo que a presença de um é suficiente para satisfazer a valência semântica de um predicador. No *frame* *Travel*, os EFs Direção, Trajetória, Destino e Origem atuam em *coreset*. Dessa forma, no exemplo dado em (10), como o EF Destino é instanciado, não é necessário que os demais também o sejam, embora isso não signifique que não possam atuar em conjunto.

- (10) [Diogo _{VIAJANTE}] VIAJOU^{ALVO} [para Bom Jardim de Minas _{DESTINO}] [para ver os tios _{RAZÃO}].

Por outro lado, a relação *exclui* impede que determinados EFs ocorram juntos. Por exemplo, os mesmos EFs Direção, Trajetória, Destino e Origem, além de atuarem em *coreset*, impedem a ocorrência do EF Área. Ou seja, se qualquer um deles estiver presente na sentença, o EF Área não será expresso, e,

consequentemente, se a Área for especificada, os demais não serão instanciados. O sintagma *pelo pantanal*, no exemplo (11), exclui a presença dos EFs supracitados.

(11) O PASSEIO^{ALVO} [pelo Pantanal_{ÁREA}] trouxe muitas surpresas.

Em contraste, a relação *requer* se dá quando um EF exige a presença de outro. Para um estar presente, o outro também deve ser anunciado. O *frame Attaching* (Anexar) expressa essa relação: em (12), para que o EF Item, *o aviso do café da manhã*, ocorresse, fez-se necessária a presença do EF Alvo, *no mural*, e vice-versa.

(12) [O recepcionista_{AGENTE}] AFIXOU^{ALVO} [o aviso do café da manhã_{ITEM}] [no mural_{ALVO}].

Além das relações entre os Elementos de *Frame*, a FrameNet também registra os casos em que os EFs nucleares não vêm expressos na sentença. Esses casos são divididos em Instanciação Nula Definida (DNI), Instanciação Nula Indefinida (INI) e Instanciação Nula Construcional (CNI). A Instanciação Nula Definida ocorre quando determinado elemento de *frame* ausente é recuperado pelo contexto linguístico. Geralmente, essa instanciação é verificada através de mecanismos de coesão referencial, como em casos de elipse e anáfora, por exemplo.

(13) VIAJAMOS^{ALVO} [para fugir da nossa rotina_{FINALIDADE}], mas somos seduzidos pela rotina dos outros e ter um lugar certo onde comprar a baguette para o café da manhã passa a ser um indispensável prazer parisiense. [VIAJANTE_{DNI}] [DESTINO_{DNI}] (Cetenfolha/Cetempúblico)

Na sentença (13), embora a elipse tenha omitido o sujeito gramatical da sentença, que faz referência ao EF Viajante, compreende-se, pela flexão verbal, que se trata de um discurso na primeira pessoa do plural, *nós*, que, certamente, será recuperado pelo contexto. Outro EF omitido nesse exemplo é o Destino, que não é expresso no escopo da oração, porém, através das informações anunciadas ao decorrer do texto, chega-se à conclusão de que se trata de Paris. Além desse tipo de ocorrência, há também aquelas cujo elemento não é recuperado lexicalmente, mas, por evocar *frames* com traços particulares, também são consideradas Instanciações Nulas Definidas.

(14) Ele consegue fazer de tudo com a perna esquerda: **chutar** forte, **cruzar**, **passar**, elogiou certa vez o ex-craque alemão Günter Netzer. (FIFA)

A combinação das ações sugeridas por *chutar*, *cruzar* e *passar*, em (14), estrutura Unidades Lexicais que evocam *frames* do domínio do futebol. Dessa forma, o sintagma que ocupará esse espaço será um referente específico, *bola*, único item possível nesse contexto. Diferentemente desses casos, há EFs que não são mencionados no corpo do texto, e, embora possam ser inferidos, não é possível dizer ao certo de quais se tratam. Ocorrências assim são consideradas tipos de Instanciação Nula Indefinida. Também nomeados como existenciais, os casos de INI são aqueles em que o referente está ausente do texto. Em Ruppenhofer et al. (2010), alguns exemplos acenam para verbos cujos comportamentos são semelhantes a *comer*, *costurar* e *beber*. Embora considerados transitivos, esses verbos admitem casos intransitivos, como em:

(15) Em época de Natal, as famílias passam o dia **cozinhando** para a ceia.

(16) João **bebeu** à noite inteira, mas chegou em casa de pé.

Contrastando com os casos de DNI, na Instanciação Nula Indefinida, os elementos ausentes não são informados pelo discurso. Ainda que haja prováveis respostas como *peru* ou *frango*, em (15), e *bebida alcoólica*, em (16), a valência assumida por esses predicadores permite a ausência de tais referentes, mas não possibilita afirmar de quais itens lexicais se tratam especificamente. Para a realização deste trabalho, essa análise foi reconsiderada no que diz respeito ao EF Turista. Em vários exemplos, esse Elemento de *Frame* não esteve anunciado no escopo do alvo e também não pôde ser recuperado lexicalmente pelo contexto. Dessa forma, a proposta do *The book* seria considerá-lo como caso de Instanciação Nula Indefinida. Contudo, pelo fato de os *corpora* serem exclusivos do domínio turístico, salvo o *corpus* da FIFA, ficou implícito que o EF Turista pode ter como referente, em última análise, o próprio leitor dos textos, já que este segue sendo o ator principal da atividade turística.

(17) [A cidade LUGAR] surpreende [pela beleza de seu conjunto arquitetônico ATRAÇÃO]. [DNI Turista] (Vai_Brasil)

(18) [A região LUGAR] oferece [uma rica e eclética culinária a base dos frutos do mar ATRAÇÃO]. [DNI Turista] (BrazilTour)

Sentenças do tipo (17) e (18) não especificam para quem se destinam essas atrações. Entretanto, entende-se que se trata de um turista na medida em que os *corpora* especificam as atividades turísticas que podem interessar os turistas. O

outro tipo de Instanciação Nula é o construcional. Nesses casos, a omissão do EF acontece devido à construção gramatical em que o alvo aparece. Por isso, os constituintes são também considerados estruturalmente omitidos.

- (19) Viaje com roupas leves e opte por calçados confortáveis.
- (20) Gramado é uma cidade visitada em todos os invernos.

Em (19) e (20), as sentenças são estruturadas de modo que o sujeito gramatical possa ficar ausente da estrutura lexical. A primeira é uma frase imperativa, e, por isso, o agente não é expreso; a segunda é passiva, tornando facultativa a presença do agente. A CNI acontece também diante de orações introduzidas por infinitivo instrucional, como em prescrições médicas.

Além da adoção desses critérios, a FrameNet também anota Tipos Semânticos. Conforme ressaltado por Ruppenhofer et al. (2010, p. 79), a função dessa anotação é registrar as informações que não estão incluídas nas hierarquias dos *frames* e dos seus elementos. Esse estudo inclui os tipos de Elementos de *Frame*, os tipos de *frames* e os tipos de Unidades Lexicais. Quanto aos tipos semânticos de EFs, há vários identificados. Em *Travel*, a Área é considerada de tipo semântico Localização, já o Viajante, Consciente.

Além dos traços semânticos, a FrameNet também tem o cuidado de descrever os dados pelo viés sintático. A soma desses princípios garante a esse banco de dados ser, de acordo com Fillmore (2006), o único projeto lexicográfico com tais características. O processo de análise de dados é conhecido como anotação em camadas. Elemento de *Frame*, Função Gramatical e Tipo Sintagmático, são as três camadas que, necessariamente, são detalhadas.

As entradas na FrameNet são construídas da seguinte forma. Um grupo de Unidades Lexicais é escolhido como representante de um dado *frame*; analisa-se uma UL por vez; sentenças exemplo são extraídas do *corpus* e classificadas pelo contexto sintático; amostras representativas são selecionadas, ilustrando claramente o sentido em questão; e essas sentenças são anotadas de acordo com o *frame* que a UL evoca. Antes de a anotação começar, etiquetas são escolhidas para representar os papéis semânticos, ou elementos do *frame* (EFs), que a UL apresenta em relação ao *frame* em questão, por exemplo, Comprador, Vendedor, Produtos, Dinheiro etc para os diversos *frames* relacionados ao Comércio. Aos constituintes das sentenças, gramaticalmente conectados à UL em questão, ou seja, aos EFs, designam-se etiquetas apropriadas; são também atribuídas a eles informações a respeito da função gramatical (FG) - sujeito, objeto etc - e tipo sintagmático (TS), por exemplo, oração subordinada finita, SV infinitivo, SN etc. (FILLMORE, 2006, p. 617)¹⁵

¹⁵ "FrameNet entries are built up in the following way. A group of LUs is chosen as representative of a single frame; taking one LU at a time; example sentences containing it are extracted from the corpus

Para realizar cada uma dessas fases, há dois padrões de anotação com os quais a FrameNet trabalha: o de texto corrido e o lexicográfico. A anotação de texto corrido objetiva tratar de todos os *frames* que forem surgindo no texto. Com esse procedimento, as Unidades Lexicais não são escolhidas pelo anotador, o intuito é que o texto vá guiando o aparecimento das ULs e, conseqüentemente, a análise dos *frames*.

45. Still, the future of **Las Vegas** is **SURE** Likelihood to be **DETERMINED** Contingency, as much by the pioneering spirit that **BUILT** Building the **CITY** Locate_by_use as by anything else. With its **NEW** Age look and **NEW** Age **AIM** Purpose, the prospects are **GOOD** Desirability despite the warnings. Inevitably, some **PEOPLE** People will get burned in the **PROCESS** Process, but **OTHERS** Increment will **RISE** Motion_directional triumphantly from the fray to even **GREATER** Desirability **SUCCESSSES** Success_or_failure. And that, more than anything, is the **ONE** Cardinal_numbers constant that **CHARACTERIZES** Distinctiveness the **PAST** Individual_history, present, and future of Las Vegas.

Annotator ID(s): 143, 225, 361, 571, 589, 597, 618, 976

Clear Sentences Turn Colors Off

Still, the future of Las Vegas is **SURE** to be determined as much by the pioneering spirit that built the city as by anything else.

Still, the future of Las Vegas is sure to be **DETERMINED** as much by the pioneering spirit that built the city as by anything else.

Still, the future of Las Vegas is sure to be determined as much by the pioneering spirit that **BUILT** the city as by anything else.

Still, the future of Las Vegas is sure to be determined as much by the pioneering spirit that built the **CITY** as by anything else.

With its **NEW** look and new **aim**, the prospects are good despite the warnings. **INI**

With its new look and **NEW** aim, the prospects are good despite the warnings. **INI**

With **its** new look and new **AIM**, the prospects are good despite the warnings. **INI**

With its new look and new aim, the prospects are **GOOD** despite the warnings.

Inevitably, some **PEOPLE** will get burned in the process, but others will rise triumphantly from the fray to even greater successes.

Inevitably, some people will get burned in the **PROCESS**, but others will rise triumphantly from the fray to even greater successes.

Inevitably, some people will get burned in the process, but **OTHERS** will rise triumphantly from the fray to even greater successes. **DNIDNI**

Inevitably, some people will get burned in the process, but **others** will **RISE** triumphantly from the fray to even greater successes.

Inevitably, some people will get burned in the process, but others will rise triumphantly from the fray to even **GREATER** successes. **INIINI**

Inevitably, some people will get burned in the process, but others will rise triumphantly from the fray to even greater **SUCCESSSES**. **INIINI**

And that, more than anything, is the **ONE** constant that characterizes the past, present, and future of Las Vegas.

Figura 19: Anotação de Texto Corrido: História de Las Vegas

Fonte: <<http://framenet.icsi.berkeley.edu/fndrupal/index.php?q=fulltextIndex>>

Como já se observou, a anotação de texto corrido propicia o conhecimento dos *frames* evocados no texto. Essa vantagem faz com que a Semântica de *Frames* seja útil para a Linguística Textual, a Análise do Discurso e também áreas como Comunicação Publicitária e *Marketing*. A promoção de produtos e ideias no mercado e na sociedade, de modo geral, faz com que especialistas desse ramo tenham de planejar maneiras eficientes para convencer o público. Para tal propósito, conhecer os *frames* que se deseja acionar, é, sem dúvida, uma boa estratégia.

Na Figura 19, o trecho 45 da História de Las Vegas, extraído do American National Corpus, é ilustrado pela FrameNet para mostrar como é o procedimento da anotação de texto corrido. No texto, as palavras destacadas em amarelo

and sorted by syntactic context; representative samples are selected that clearly illustrate the sense in question; and the selected sentences are annotated according to the frame the LU evokes. Before the annotation begins, labels are chosen to represent the semantic roles or frame elements (FEs) that the LU has in respect to the given frame, for example, Buyer, Seller, Goods, Money, etc., for the various frames connected with Commerce. Sentence constituents that are grammatically linked to the LU in question are assigned appropriate FE labels, and the constituents thus labeled are also provided with information about their grammatical function (GF) – Subject, Object, etc. – and their phrase type (PT), e.g., finite *that*-clause, marked infinitive VP, NP, etc”

representam as entidades mencionadas, como o nome próprio *Las Vegas*; em itálico, estão aquelas consideradas não evocadoras de *frames*, como artigos e preposições; as palavras destacadas em azul, em hiperlinks, são as unidades evocadoras de *frames*; clicando na palavra, o usuário é direcionado para a análise daquela UL; abaixo do trecho, ainda na Figura 19, estão as sentenças analisadas separadamente a partir dos hiperlinks acionados.

Já a anotação lexicográfica tem como finalidade selecionar apenas uma Unidade Lexical para registrar a valência semântica e sintática de seus constituintes dentro de uma sentença. Esse é o procedimento tradicional e mais desenvolvido pela FrameNet. Na estruturação de dicionário, é o mais adequado, pois parte de um sentido específico para explorar as possibilidades semânticas e sintáticas dos constituintes que acompanham a unidade alvo.

Assim como acontece no processo de anotação de texto corrido, a anotação lexicográfica também respeita a localidade sintática (RUPPENHOFFER ET AL., 2010, p. 27). Esse princípio também conhecido como projeção máxima da Unidade Lexical prevê que todos os Elementos de *Frame* que devem ser analisados estejam localizados dentro da projeção máxima nucleada pela UL. Os sujeitos sintáticos têm um tratamento diferenciado, isentando-se dessa regra. Geralmente, isso acontece em casos de alçamento e controle e com sujeitos inseridos por sentenças relativas. A título de exemplo, em (21), *Thiago Luís* é argumento externo do verbo *querer* e também do verbo *viajar*, recebendo papel temático de cada um deles, diz-se, nesse caso, ocorrer controle de sujeito sobre sujeito. Já o sintagma *Thiago Luís*, em (22), atua como argumento interno do verbo *convencer* e como argumento externo do verbo *viajar*, daí ser reconhecido como controle de objeto sobre sujeito.

(21) [Thiago Luís VIAJANTE] quer VIAJAR^{ALVO} [para a praia DESTINO] [nas férias TEMPO].

(22) Ana Carolina convenceu [Thiago Luís VIAJANTE] a VIAJAR^{ALVO} [com ela ACOMPANHANTE] [para Olímpia DESTINO] [nas férias TEMPO].

Diferentemente de situações desse tipo, cujos verbos selecionam o mesmo referente, reconhecido através de uma relação anafórica, há casos cujos argumentos de determinado verbo se movem para uma posição atemática. *Ana Carolina*, em (23), embora atue como sujeito do verbo *dever*, não é argumento dele, mas sim de *viajar*. Esse caso é denominado alçamento de sujeito para sujeito. Na

sentença (24), *Thiago Luís* é argumento interno do verbo *fazer*, mas atua como externo do verbo *viajar*, esse é outro caso de alçamento, de sujeito para objeto.

- (23) [Ana Carolina VIAJANTE] deve VIAJAR^{ALVO} [com Thiago Luís ACOMPANHANTE] [nas férias TEMPO]. [ÁREA INI]
- (24) Nós fizemos [Thiago Luís VIAJANTE] VIAJAR^{ALVO} [com Ana Carolina ACOMPANHANTE]. [ÁREA INI]

Além desses casos, há também aqueles estruturados com sentenças relativas. Nesses, tanto o elemento de *frame* não local quanto o pronome relativo são marcados e anotados. No exemplo ilustrado pela sentença (25), tanto o sintagma *famílias* como o pronome relativo *que* representam o EF Viajante. Essa ocorrência é anotada em outra camada, chamada *Outros*, onde se assinala a existência do Antecedente (Ant), *famílias*, e do Relativo (Rel), *que*.

- (25) [Famílias VIAJANTE] [que VIAJANTE] VIAJAM^{ALVO} [para a cidade maravilhosa DESTINO] [nesse feriado TEMPO] devem preferir os horários da manhã.

Assim como ilustrado pela Figura 13, que representou, a partir de uma análise lexicográfica, o resultado por *Travel* na opção *Annotation* em *Lexical Unit Index*, a Figura 20 demonstra esse mesmo processo. A diferença é que os dados no *software* de anotação exibem os EFs com etiquetagem em cores, por isso, o primeiro passo é definir quais são os Elementos de *Frame*, atribuindo uma etiqueta a cada um. Atendendo ao princípio da localidade sintática, os constituintes etiquetados se encontram gramaticalmente vinculados à Unidade Lexical, que é destacada em preto. Cada cor corresponde a um elemento de *frame* específico, que é definido na primeira camada de anotação.

429-s20-rcoll-aircraft	1. Her Royal Highness TRAVELLED in an aircraft of The Queen's Flight INI
429-s20-rcoll-country	1. If you TRAVEL to a tropical country, write to that country's tourist office to tell them that you are concerned about the preservation of their forests . 2. Paul Pratt TRAVELLED through forty-eight countries . 3. He has TRAVELLED with the team to every country in Europe and his outstanding memory was Ireland 's victory in the European Helvetia Cup in Norway in 1981 . 4. Customs officers from each country would be allowed to TRAVEL to either country in pursuit of suspected traffickers 5. The Essex Section will TRAVEL by coach to the poppy country of Picardy , Northern France to study public transport and railway installations and services in Amiens and district .
429-s20-rcoll-world	1. Although he has TRAVELLED the world , Edinburgh and the family memories it evokes , are still profoundly important to Graeme Souness . 2. The owner , who had recently died , had TRAVELLED the world . 3. Carefully planning the cheapest air fares , allocating her meagre budget as sparingly as possible on food and accommodation , she TRAVELLED the world and took on the best . 4. She TRAVELLED round the world , paid two visits to the United States , and for a time taught English in Japan . 5. He 'll read extracts from the works of American writers who TRAVELLED to Paris after the first world war .
680-s20-pass	1. Days began early and ended late so that maximum distances could be TRAVELLED .CNI
710-s20-adverbabroad	1. After a five-year apprenticeship to Richard Lane of Manchester , he TRAVELLED abroad for a year and sketched the most extraordinary number of buildings and architectural details . 2. From 1802 to 1810 he TRAVELLED abroad, to Berlin , Dresden , Naples , Rome , Milan , and Vienna, partly as a representative of his firm . 3. The two girls who shared the upstairs flat were models , and often TRAVELLED abroad on assignments . 4. It meant that she could see you every day , work closely with you , even TRAVEL abroad with you ." 5. They can TRAVEL abroad going to shows and fights .
710-s20-adverbtogether	1. We could TRAVEL together as far as Doha and I would go on alone to Abu Dhabi . 2. We were going to TRAVEL together as far as Paris , I was to go on to Calais .
730-s20-ppacross	1. When he TRAVELLED across America, it was not so much the places as the travelling which found its way into his pictures . 2. Late in 1981 he TRAVELLED across northern China , from Peking in the east to Xian , Urumchi , Turpan and Kashgar in the west . 3. He had TRAVELLED across the city from the suburbs to the apartments of the ruling elite . 4. By 1907 , she was the madam of the most luxurious of the Dublin brothels , and she had TRAVELLED extensively across Europe .

Figura 20: Anotação lexicográfica: sentenças que ilustram a UL *viajar* no *frame* *Travel*
Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=lulnde>>

Assim, na Figura 21, que ilustra a sentença *Menores de idade podem viajar sozinhos ou desacompanhados dos pais ou responsáveis*, o sintagma *Menores de idade* corresponde ao EF Viajante; o sintagma *sozinhos ou desacompanhados dos pais e responsáveis*, ao EF Maneira. Como os EFs nucleares Direção, Destino, Trajetória e Origem atuam em *Coreset*, apenas o Destino foi marcado como INI.

Layer	Menores de idade podem viajar sozinhos ou desacompanhados																			
FE	Viajante										Maneira									
GF_br	Ext										Dep									
PT_br	NP										AdjP									
Outr...																				

	Viagem.viajar.v - FE	Viagem.viajar.v - GF_br	Viagem.viajar.v - PT_br	Viagem.viajar.v - Outros_br
● Destino <D>	INI			○ Lugar <P>
● Direção <F6>				○ Maneira <M>
● Meio_de_tra...				○ Meio <F12>
● Origem <S>				○ Tempo <F2>
● Trajetória <...>				○ Velocidade ...

Figura 21: Anotação da Unidade Lexical *viajar* na FrameNet Brasil

Algumas situações merecem atenção especial durante a etapa de etiquetagem dos Elementos de *Frame*. Há casos, por exemplo, em que um sintagma designa mais de um EF, esse fenômeno é chamado de *conflação*. Na Figura 22, o sintagma *os atrativos marketing do Jalapão* agrega o EF Atração e o EF Lugar. A prática adotada nesses casos é a duplicação da camada EF: no caso, uma é especificada como Atração, e a outra, como Lugar.

Layer	pra visitar os atrativos marketing do Jalapão																			
FE	Atração										Lugar									
GF_br	ObjD										ObjD									
PT_br	NP										NP									
Outros...																				

Figura 22: Anotação de conflação de Elemento de *Frame*

Outro caso também previsto pela FrameNet é a incorporação de Elemento de *Frame* pela Unidade Lexical. Na Figura 23, *atrações* é a Unidade Lexical que evoca o *frame* *Turismo_de_atração*, entretanto também designa o EF Atração. Por esse motivo, a UL é marcada como INC (Incorporation), indicando a incorporação desse EF pela UL.

Layer	P o s s u i , t a m b é m , a l g u m a s a t r a ç õ e s h i s t ó r i c a s																												
FE											Q	u	a	n	t	i	d					T	i	p	o				
GF_br											Q	u	a	n	t					D	e	p							
PT_br											N	u	m					A	d	J	P								
Outr...																													

Turismo_de_atração.atração.n - FE		Turismo_de_atração.atração.n - GF_br		Turismo_de_atraçã...	
<input checked="" type="radio"/> Atração <F2>	INC	<input type="radio"/> Meio <>			
<input checked="" type="radio"/> Turista <F1>	DNI	<input type="radio"/> Origem <>			
<input type="radio"/> Duração <>		<input type="radio"/> Quantidade ...			
<input type="radio"/> Finalidade <...>		<input type="radio"/> Razão <>			
<input type="radio"/> Frequência <>		<input type="radio"/> Tempo <>			
<input type="radio"/> Lugar <F3>		<input type="radio"/> Tipo <>			
<input type="radio"/> Maneira <M>		<input checked="" type="radio"/> Acompanha...			

Figura 23: Anotação de incorporação de Elemento de *Frame*

Tendo estabelecido os nomes dos EFs, inclusive marcando aqueles de instanciação nula, a segunda camada se refere à Função Gramatical. A FrameNet Brasil reconhece as seguintes funções¹⁶: Aposto, Externo (Ext), Determinante possessivo (DetPoss), Objeto Direto (ObjD), Objeto Indireto (ObjInd), Dependente (Dep), Quantificador (Quant) e Núcleo Nominal (Nucl). Aposto é a palavra ou expressão utilizada para identificar, explicar ou resumir um item nominal que esteja sendo anotado como alvo; Ext é o termo que ocupa função de argumento externo; DetPoss, geralmente, acompanha o nome, estabelecendo referência com teor de posse; ObjD designa os EFs avaliados como argumento interno, não regidos por preposição; ObjInd marca o argumento interno regido por preposição; a função Dep destaca o adjunto; Quant assinala a presença de item lexical ou construção que traz acepção de quantificação e Nucl evidencia um núcleo nominal geralmente modificado por adjetivo atributivo. A terceira camada de análise é o Tipo Sintagmático (TS), que distingue os diferentes tipos de sintagmas, como sintagma nominal, sintagma preposicionado, sintagma adverbial etc.

¹⁶ A FrameNet concebe a existência de três Funções Gramaticais: Externo, Objeto e Dependente. Na função Externo, estarão os argumentos externos, na Objeto, os argumentos internos não preposicionados e, na função Dependente, estarão tanto os argumentos internos preposicionados como os adjuntos.

Layer	N	o	s	s	o	s	t	é	c	n	i	c	o	s	v	i	a	j	a	r	a	m	m	a	i	s	d	e	3	0	m	i	l	q	u	i	l	ô	m	e	t	r	o	s						
FE	V	i	a	j	a	n	t	e														D	i	s	t	â	n	c	i	a																				
GF_br	E	x	t																			D	e	p																										
PT_br	N	P																				N	P																											
Outros...																																																		

Figura 24: Anotação semântica e sintática com a UL *viajar*

Na sentença *Nossos técnicos viajaram mais de 30 mil quilômetros*, Figura 24, o sintagma *nossos técnicos* é marcado como EF *Viajante* na camada Elemento de *Frame*; Externo, para Função Gramatical; e Sintagma Nominal para Tipo Sintagmático. Em seguida, é destacada a Unidade Lexical, *viajaram*, diante do Elemento de *Frame* *Distância, mais de 30 mil quilômetros*, que assume Função Gramatical Dep e Sintagma Nominal no Tipo Sintagmático.

Essas três camadas representam as valências semântica e sintática dos EFs. Entretanto, outras camadas são introduzidas à medida que a ocorrência de outros fenômenos vá sendo notada. A sentença, na Figura 25, é estruturada com o verbo suporte *fazer*. Nesse tipo de construção, o verbo é classificado com carga semântica fraca; quem irá se encarregar de particularizar o evento será seu complemento, *tour*, que deixará de atuar como complemento do verbo para inserir a predicação necessária como Unidade Alvo. Em casos como esse, insere-se a camada *Nome* para evidenciar a ocorrência de verbo suporte (Sup). A inclusão de camadas secundárias desse tipo decorre, principalmente, das especificidades originadas pela Unidade Lexical (UL).

Layer	F	a	ç	a	u	m	t	o	u	r	a	p	é	p	e	l	o	C	e	n	t	r	o	H	i	s	t	ó	r	i	c	o	,	p	e	l	a	s	r	u	a	s	c	a	l	ç	a	d	a	s				
FE											M	e	i	o	L	u	g	a	r																																			
Sent																																																						
GF_br																																																						
PT_br																																																						
Outr...																																																						
N_br	S	u	p	o																																																		

Figura 25: Anotação de verbo suporte

Acerca das Unidades Lexicais, é importante salientar que, ainda que sejam centrais para a estruturação dos *frames*, há também aqueles que não apresentam uma contraparte lexical. Geralmente, são *frames* que se estruturam em cenários. De modo geral, esses *frames* têm como função orientar a sistematização de *frames* produtivos, sendo um exemplo *Employment_scenario* (Cenário_do_emprego). Além de guiar a estruturação dos demais *frames*, ele também organiza as relações

que acontecem nesse cenário. A Figura 26 traz sua visualização juntamente com a Figura 27, que informa a legenda com os respectivos nomes das relações.

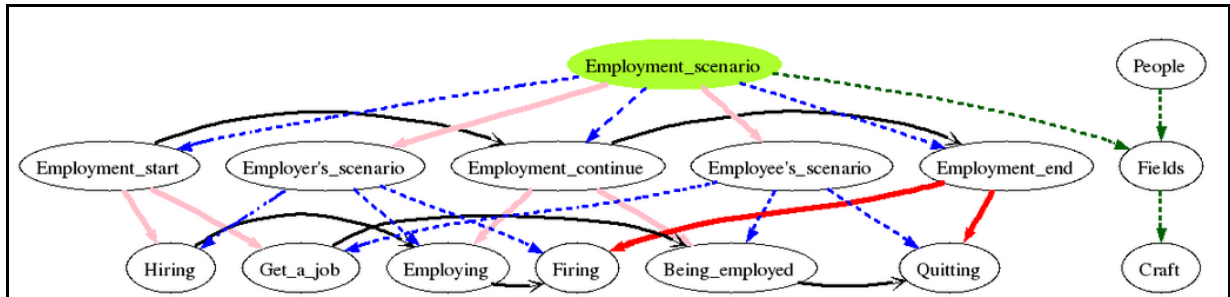


Figura 26: *Frame* Employment_scenario
 Fonte: <<https://framenet.icsi.berkeley.edu>>

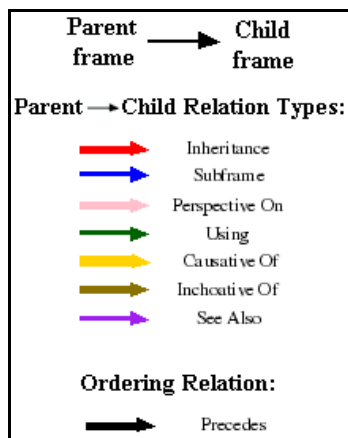


Figura 27: Legenda de relações entre *frames*
 Fonte: <<https://framenet.icsi.berkeley.edu>>

Outro tipo semântico de *frame* é o não perspectivado. Ele aborda determinado conceito de forma mais ampla, geralmente, não perfila perspectiva específica, ou trata de várias sem o intuito de detalhá-las. Um exemplo é o *Commerce_scenario* (Cenário_do_comércio). Esse *frame* inclui EFs como *Comprador*, *Dinheiro*, *Vendedor* e *Mercadoria*. Porém, não há uma perspectiva específica assumida.

- (26) Os PREÇOS^{ALVO} aumentam no início das estações [INI DINHEIRO] [INI COMPRADOR] [INI MERCADORIA] [INI VENDEDOR].
- (27) Nos finais de ano, triplica o movimento do COMÉRCIO^{ALVO} [INI DINHEIRO] [INI COMPRADOR] [INI MERCADORIA] [INI VENDEDOR].

Nas sentenças (26) e (27), as Unidades Lexicais *preços* e *comércio* evocam o *frame* *Commerce_scenario*. O evento comercial é abordado de maneira generalizada, prova disso são as instanciações nulas encontradas nessas sentenças, que reforçam o caráter não perspectivado do *frame*. Essa situação fica evidenciada na sentença (28), em que o Elemento de *Frame* *Comprador* é

perfilado, fazendo com que a perspectiva assumida seja da compra. Dessa forma, o *frame* em questão será `Commerce_buy` (Comércio_compra)

(28) [Nos finais de ano TEMPO], [milhares de trabalhadores COMPRADOR] COMPRAM
[presentes MERCADORIA] [para toda a família BENEFICIÁRIO]. [INI DINHEIRO] [INI VENDEDOR].

Outra vantagem da metodologia da FrameNet é o detalhamento em torno das relações entre os *frames*. O que a autentica como uma rede de *frames* – Frame+Net – é o fato de especificar como eventos, situações e experiências mantêm relações entre si. Para sinalizar esse comportamento bem como sistematizar o contato entre *frames*, a FrameNet estabelece a rede de relações entre *frames*. São elas: *Inheritance* (Herança), *Using* (Uso), *Subframe* (Subframe), *Perspective_on* (Perspectiva), *Precedes* (Precedência) e *Causative_of / Inchoative_of* (Causativo_de / Inchoativo_de).

A relação de Herança, reconhecida também como *subtipo de*, dá-se quando um *frame* mais amplo gera outros mais específicos. Fillmore et al. (2003) destacam que, nessa relação, todos os Elementos de *Frame*, subframes e tipos semânticos do *frame* mãe terão um corresponde igual ou semelhante no *frame* filho. Na Figura 26, *Firing* (Demitir) herda do *frame* `Employment_end` (Final_do_emprego). Nesse sentido, os EFs *Employee* (Empregador) e *Employer* (Empregado) presentes no *frame* mãe se encontrarão também presentes no filho.

Uma relação que compartilha traços com a de Herança é a de Uso. Às vezes, um *frame* faz referência a outro geralmente mais abstrato. `Path_traveled` (Caminho_percorrido) estabelece relação de uso com o *frame* `Motion` (Movimento). Para que se compreenda o primeiro, é necessário ter o segundo como *background*. Ou seja, não se pode estruturar um *frame* que trata de caminho percorrido sem ter como conhecimento prévio a noção de movimento. Nessa relação, não é necessário haver as correspondências requeridas pela relação Herança.

A relação Subframe é outra com afinidade à relação Herança. Nela, há também um *frame* filho e um *frame* mãe, entretanto, aquele é um subevento de um evento mais complexo, representado pelo *frame* mãe. Geralmente, esses subframes se referem a sequências de estados de coisas ou sequências temporais e apresentam particularidades suficientes para serem descritos em *frames* diferentes. Ainda que os EFs possam se dirigir, de alguma forma, ao *frame* mãe, isso não se constitui como regra.

Quando essa ordenação de subeventos é explicada de modo temporal, isto é, em etapas, os *frames* se relacionam através da relação Precedência. Os cenários *Employment_start* (Início_do_emprego), *Employment_continue* (Emprego_em_andamento) e *Employment_end* (Término_do_emprego) são subframes que sinalizam uma ordenação temporal em *Employment_scenario*. Os acontecimentos tendem a seguir uma ordem cronológica: o início de um emprego, a atuação do funcionário em determinado cargo e, por fim, o término da atividade, quando tem fim o contrato entre as partes.

Outra relação prevista pela FN é a de Perspectiva. Nessa, ainda que os *frames* compartilhem um contexto em comum, eles focalizam aspectos diferentes. O *Employer's_scenario* (Cenário_do_empregador) reflete a perspectiva daquele que fornece a vaga de emprego e *Employee's_scenario* (Cenário_do_empregado) perfila o funcionário; ambos são perspectivas do *frame* *Employment_scenario*.

Há também a indicação de outras duas relações verificadas entre *frames* estativos, causativos e incoativos¹⁷, que são *Incoativo_de / Causativo_de*. Uma exemplificação desses casos pode ser dada com os *frames* *Cause_change_of_position_on_a_scale* (Causar_mudança_de_posição_em_escala), que é causativo de *Change_position_on_a_scale* (Mudar_posição_em_escala), que, por sua vez, é incoativo de *Position_on_a_scale* (Posição_em_escala).

Na sentença (29), o *frame* evocado é o causativo *Causar_mudança_de_posição_em_escala*, o sujeito, os *turistas*, atua como agente causador da ação; em (30), trata-se de *Mudar_posição_em_escala*, o *frame* é incoativo, na medida em que o sujeito sintático, o EF Atributo, o *índice de stress*, não é o responsável pela ação; e, em (31), suscita-se o *frame* *Posição_em_escala*, por se tratar de um evento estativo, cujo sujeito da oração é um experienciador do estado expresso pela UL *alto*.

(29) [Os turistas _{AGENTE}] AUMENTAM ^{ALVO} [a quantidade de lixo _{ITEM}] [nas praias _{LOCAL}].

(30) [O índice de stress _{ATRIBUTO}] DIMINUI ^{ALVO} [com as férias]. [_{ITEM DNI}] [_{DIFERENÇA INI}] [_{VARIAÇÃO_DE_VALOR INI}]

(31) [Os valores de pacotes para Foz do Iguaçu _{ITEM}] são ALTOS ^{ALVO} [nessa época do ano _{DOMÍNIO}].

¹⁷ *Frames* considerados causativos são herdeiros do *frame* *Transitive_action* (Ação_transitiva), os incoativos, do Event (Evento), e os estativos podem herdar tanto do *frame* *Event* (Estado) quanto do *Gradable_Attribute* (Atributo_gradativo) (cf. RUPPEHOFER ET AL., 2010).

Como visto, além de mapear o comportamento semântico, a FN também descreve, baseando-se em *corpus*, a valência sintática de cada Unidade Lexical. O histórico de pesquisa de Fillmore até chegar à Semântica de *Frames* mostra o interesse em conhecer o processo de estruturação formal da linguagem. A grande questão percebida por ele nesse processo foi compreender que a esquematização das cenas em que as palavras se inseriam possibilitava conhecer suas propriedades semânticas e sintáticas.

O procedimento que reconhece os padrões de valência garante essa investigação. As propriedades sintáticas disponibilizam informações sobre os tipos sintagmáticos (sintagma nominal, preposicional etc) e as funções gramaticais (argumento externo, objeto direto e dependente) dos EFs associados às Unidades Lexicais; e a valência semântica, por sua vez, indica informações sobre quais são os elementos que ocupam essas posições.

Os Quadros 2 e 3 apresentam o sumariamento dos Elementos de *Frame* da UL. No Quadro 2, é exibida a realização de todos os Elementos de *Frame*, informando a quantidade de sentenças descritas com cada função gramatical. Como há a marcação das instanciações nulas para os EFs nucleares, é possível verificar, na segunda coluna, o número de sentenças analisadas com esse EF, no caso de *Employee* (Empregado) foram 72, sendo que, em 64 delas, sua atuação se deu como Sintagma Nominal com função de Objeto, e 8 sentenças foram com função de Externo.

Frame Element	Number Annotated	Realization(s)
Compensation	(1)	PP[on].Dep (1)
Contract basis	(1)	2nd.-- (1) NP.Dep (1) PP[on].Dep (4) Sub.Dep (1)
Employee	(72)	NP.Obj (64) NP.Ext (8)
Employer	(73)	NP.Ext (65) CNI.-- (2) PP[by].Dep (6) NP.Obj (1)
Manner	(2)	AVP.Dep (2)
Place	(10)	PP[for].Dep (1) PP[in].Dep (2) PP[at].Dep (3) PP[on].Dep (1)

Quadro 2: Realizações de Elementos de *Frame* na UL *employ*

Fonte: <<https://framenet2.icsi.berkeley.edu>>

No Quadro 3, especifica-se como se deu a combinação desses EFs nas sentenças analisadas. A inserção em cada padrão leva em conta o comportamento dos sintagmas que acompanham a unidade alvo em três critérios: EF, TS e FG. Para a UL *employ*, foram encontrados seis padrões sintáticos.

Number Annotated	Patterns				
1 TOTAL	Compensation	Employee	Employer	Employer	Position
(1)	PP[on] Dep	NP Obj	NP Ext	NP Ext	INI --
1 TOTAL	Contract basis	Employee	Employer	Manner	Task
(1)	PP[on] Dep	NP Obj	NP Ext	AVP Dep	VPto Dep
1 TOTAL	Contract basis	Employee	Employer	Place	Position
(1)	PP[on] Dep	NP Ext	CNI --	PP[for] Dep	2nd --
1 TOTAL	Contract basis	Employee	Employer	Position	Task
(1)	2nd --	NP Obj	NP Ext	2nd --	2nd --
4 TOTAL	Contract basis	Employee	Employer	Task	
(1)	NP Dep	NP Obj	NP Ext	VPto Dep	
(1)	PP[on] Dep	NP Obj	NP Ext	DNI --	
(1)	PP[on] Dep	NP Obj	NP Ext	INI --	
(1)	Sub Dep	NP Obj	NP Ext	VPto Dep	
1 TOTAL	Employee	Employer	Employer	Place	
(1)	NP Obj	NP Ext	NP Ext	PP[in] Dep	

Quadro 3: Padrões sintáticos dos Elementos de Frame na UL *employ*

Fonte: <<https://framenet2.icsi.berkeley.edu>>

O processo de anotação na FrameNet é considerado semiautomático. O *software Desktop* realiza a estatística das valências, facilitando o processo de análise de dados, mas é o anotador quem decide quais etiquetas de análise serão atribuídas a cada sintagma. É possível dizer que a FN apresenta, especialmente, um viés sintagmático para a descrição da língua, na medida em que um estudo do comportamento dos constituintes em torno da unidade alvo é de grande importância para a constituição de *frames*. Todavia, não é possível afirmar que o eixo seguido seja estritamente sintagmático. Há evidências que indicam o compromisso em tratar do eixo paradigmático: o cuidado em expor vários exemplos anotados de uma mesma Unidade Lexical, indicando as possibilidades de combinação; a listagem de várias ULs, sendo que algumas podem funcionar como substitutas umas das outras; além da relação Herança, próxima da hiperonímia / hiponímia, que pode gerar paráfrases mais genéricas a partir de *frames* mais específicos.

Ao reafirmar a teoria da Linguística Cognitiva, Talmy (2000) ressalta que as generalizações formais acerca da linguagem se dão a partir da perspectiva conceptual da língua. Nesse sentido, é válido destacar o comprometimento da FrameNet em seguir um viés cognitivista para a descrição da língua, já que a semântica é quem guia o processo de análise de dados. Como pondera Fauconnier (1997), a forma não apresenta significado, mas ela escolhe as regularidades que funcionam durante todo o processo de significação. Daí a importância do estudo sintático, pois, como as regularidades estruturais são evidenciadas, é possível construir generalizações acerca dos elementos que acompanham o *frame*, o que é útil para tarefas de PLN.

Fillmore, ao fazer uma análise da pesquisa na FrameNet, destaca que o trabalho em Semântica de *Frames* apresenta uma parte penosa e uma parte divertida: a parte que requer atenção meticulosa é o trabalho na construção de um léxico baseado em *frames* e o objetivo de integrá-lo na descrição da gramática da língua; o *divertido* é analisar como as próprias escolhas pelos itens lexicais evocam *frames* e, conseqüentemente, os resultados de tais escolhas. (FILLMORE, 2006, p. 620).

2.2.3 Expansão da FrameNet para Outras Línguas

O recurso lexicográfico em inglês disponibilizado pela FrameNet tem sido desenvolvido para outras línguas do mundo. Pesquisadores chineses, espanhóis, alemães, japoneses, suecos e também brasileiros estão construindo contrapartes dessa rede semântica para suas respectivas línguas.

A Chinese FrameNet é uma base de dados que já conta com cerca de 300 *frames*, mais de 3.000 ULs ilustradas em cerca de 18.300 sentenças anotadas com as informações semânticas e sintáticas, conforme prevê a FN. Além de procurar cobrir a gama de informações dos diversos domínios da língua, há também dados de áreas específicas, como turismo e direito.

Para o alemão, há, atualmente, três grupos de pesquisa que atuam em colaboração, o maior deles é o Projeto SALSA¹⁸ em Saarbrücken. Em Stuttgart, a

¹⁸ Para mais informações do projeto SALSA: <<http://www.coli.uni-saarland.de/projects/salsa/page.php?id=index>>

equipe vem trabalhando em técnicas de extração de informação. Em Austin, o grupo¹⁹ liderado pelo pesquisador Hans C. Boas utiliza as informações disponibilizadas pelo SALSA para detalhar os dados da contraparte alemã da FrameNet e desenvolver ferramentas com tal suporte teórico.

A Spanish FrameNet²⁰, sediada na Universidade Autônoma de Barcelona, inclui pesquisadores de várias universidades espanholas, dentre eles o professor Carlos Subirats. O *corpus* construído conta com mais de 300 milhões de palavras. A sentença (32), da base de dados da FN do espanhol, ilustra uma ocorrência com a Unidade Lexical *llegar* (*chegar*).

- (32) [El presidente uruguayo, Julio María Sanguinetti _{THEME}] LLEGÓ^{TARGET} [hoy _{TIME}] [para una visita oficial de tres días _{PURPOSE}] y mañana firmará con su colega paraguayo, Juan Carlos Wasmosy, dos convenios de cooperación agropecuaria y judicial. [_{DNI} Goal]

No exemplo, o EF nuclear *Theme* (Tema), que se refere àquele que se move, é representado por *El presidente uruguayo, Julio María Sanguinetti*; o EF *Goal* (Destino) aparece como Instanciação Nula Definida, os demais instanciados na sentença, EFs *Time* (Tempo) e *Purpose* (Finalidade), tratam-se de EFs não nucleares, vide outros exemplos na Figura 28.

- V-A-(1)
 1. Los contratos suponen grandes inversiones para el club y es conveniente evitar la posibilidad , por remota que sea , de que [_{Theme}un jugador] pueda LLEGAR^{TARGET} [_{Goal}DNI]
- V-GN-(1)
 1. Dentro de las investigaciones , los agentes investigadores acudieron a l domicilio de Humberto Barrera Guerra , porque presuntamente está a su nombre [_{Mode_of} [Mode_of_transportation]en que] LLEGÓ^{TARGET} [_{Theme}la víctima] . [_{Goal}DNI]
 2. [_{Theme}Jiang y otros altos funcionarios] LLEGARON^{TARGET} [_{Purpose}para inspeccionar las medidas tomadas para luchar contra las inundaciones en la provincia de] agencia oficial Xinlma . [_{Goal}DNI]
- V-Prep-(1)
 1. Custodiado de día y noche por miembros de organismos internacionales , el poblado continúa su vida cotidiana , y lejos de lo que se pudiera pensar , [_{Theme}indígenas] siguen LLEGANDO^{TARGET} [_{Goal}ahí] [_{Purpose}para acompañar los en su tragedia] .
 2. [_{Theme}El presidente uruguayo , Julio Maria Sanguinetti ,] LLEGÓ^{TARGET} [_{Time}hoy] [_{Purpose}Para una visita oficial de tres días] y mañana firmará con su colega para dos convenios de cooperación agropecuaria y judicial . [_{Goal}DNI]
- V-Adv-(1)
 1. [_{Theme}El cantante y actor de 23 años de edad] , [_{Theme}que] LLEGÓ^{TARGET} [_{Time}anoche] [_{Goal}aquí] [_{Source}procedente de México] , dijo que comenzará en breve la grabación de un piloto de una serie titulada « Barefoot in Paradise » para la cadena televisiva ABC .
 2. [_{Theme}Los presidentes centroamericanos] LLEGARON^{TARGET} [_{Time}hoy] [_{Goal}aquí] [_{Purpose}para despedir a la presidenta Violeta Barrios de Chamorro y discutir la integración económica regional y la deuda externa de Nicaragua con la región] .

Figura 28: Unidade Lexical *llegar* na base de dados da Spanish FrameNet

Fonte: <<http://sfnlinux1.uab.es:8080/farina-web//LUIndex.html>>

O Japão é outro país que desenvolve uma rede semântica baseada em *frames*: a Japanese FrameNet.²¹ Desenvolvida pela Universidade de Keio com apoio de outras universidades japonesas e fomento do Ministério da Educação, Cultura,

¹⁹ Para mais informações da German FrameNet: <<http://www.laits.utexas.edu/gframenet/>>

²⁰ Para mais informações da Spanish FrameNet: <<http://sfn.uab.es:8080/SFN>>

²¹ Para mais informações da Japanese FrameNet: <<http://jfn.st.hc.keio.ac.jp/>>

Esporte e Ciência e Tecnologia do país, a equipe vê a FrameNet como uma forma de superar as barreiras impostas pela língua. Em entrevista, a pesquisadora Kyoko Ohara explana questões importantes sobre as quais esses pesquisadores vêm se dedicando. Como ela pontua, ensinar os computadores a entender a nossa linguagem passa primeiro por compreender as “questões chave” acerca da nossa capacidade de processamento.

Quando você procura em um dicionário, o que você encontra são definições como o oposto de *esquerda* é *direita*, e *leste* é o oposto de *oeste*. Porém, entender o significado de cada palavra envolve não apenas um conhecimento puramente linguístico do significado de palavras como essas, mas também conhecimento enciclopédico. Assim, nas nossas descrições dos significados das palavras, queremos incorporar tais conhecimentos enciclopédicos, incluindo o senso comum e o saber científico, o qual não é geralmente encontrado em dicionários. Esse é o objetivo do nosso projeto.²²

Nesse sentido, a professora ressalta o papel central do uso de corpus para esse propósito, já que, para analisar como os japoneses usam a língua e atribuem significado a ela, é necessário ter contato com dados que mostrem como eles *leem, falam e escrevem*.

A Suécia também é um país que desenvolve um banco de dados lexicais de acordo com os moldes da FrameNet. Pesquisadores da Universidade de Gotemburgo estão estruturando a Swedish FrameNet.²³ O projeto objetiva construir um recurso com cobertura de pelo menos 50.000 Unidades Lexicais, que contenha a descrição semântica e sintática dos dados.

No Brasil, o recurso vem sendo desenvolvido em Minas Gerais, na Universidade Federal de Juiz de Fora. A equipe tem se empenhado na construção de *corpora* representativos para viabilizar a descrição do português brasileiro pelos moldes da FrameNet. Atualmente, a FrameNet Brasil²⁴ está atuando em dois subprojetos: *Frames* e *Construções e Copa 2014*. O primeiro anota construções gramaticais com o intuito de inserir o *Constructicon* na FN Br, e o segundo, ao qual a

²² Tradução disponibilizada para inglês da entrevista com Kyoko Ohara: “When you look up a dictionary, what you find is definitions such as *left* is the opposite of *right*, and *east* is the opposite of *west*. But to understand the meaning of each word involves not just knowing purely linguistic meaning of the word like that, but also having encyclopedic knowledge of it. So in our descriptions of the meaning of words, we want to incorporate such encyclopedic knowledge, including common sense and scientific knowledge, which is not usually found in dictionaries. That’s the aim of our project”. Disponível em: <<http://www.youtube.com/watch?v=kfqR9aUcp1c>>. Acessado em 10 de outubro de 2012.

²³ Para mais informações da Swedish FrameNet: <<http://spraakbanken.gu.se/eng/swefn>>

²⁴ Para mais informações da FrameNet Brasil: <<http://www.framenetbr.ufjf.br/>>

presente dissertação se vincula, volta-se para o desenvolvimento de um dicionário eletrônico trilingue para a Copa do Mundo, que será realizada no Brasil em 2014.

No processo de construção de redes semânticas similares para outras línguas, é de se realçar a importância da adoção dos mesmos critérios de anotação da FrameNet, marcando, sempre que necessário, as particularidades de suas línguas. Esse tem sido o cuidado da FrameNet Brasil ao incluir funções gramaticais além daquelas previstas pela FN; quando avalia, de forma contrastiva, o comportamento dos Elementos de *Frame* entre o inglês e o português, fazendo as devidas modificações quando preciso. Esse procedimento permitirá construir iniciativas multilíngues que sejam capazes de conectar as framenets. Daí a necessidade de hoje esses bancos de dados manterem contato com a plataforma mãe. Propósito semelhante é compartilhado pela WordNet, banco de dados discutido na seção seguinte.

2.3 A WordNet

A WordNet,²⁵ doravante WN, é uma base de dados lexicais, originalmente construída nos Estados Unidos pela Universidade de Princeton para representar o léxico da língua inglesa americana. Os primeiros estudos, datados da década de oitenta, denotaram tentativas de entender como crianças aprendiam novas palavras. O objetivo era simular esse processo para que fosse possível facilitá-lo. Entretanto, conforme avaliaram seus precursores, essa pesquisa acabou trazendo importantes descobertas sobre a relação estabelecida entre as palavras. (MILLER, 1993, 1995; FELLBAUM, 1998; FELLBAUM & MILLER, 2007).

O objetivo desta seção é tratar de tais avanços alcançados a partir da estruturação dessa ferramenta. Para tanto, pretende-se abordar como a WN organiza as relações entre as palavras e, em seguida, seu desenvolvimento em outras línguas bem como a apresentação de aplicativos baseados nesse banco de dados.

2.3.1 O Léxico na WordNet

²⁵ Quando a palavra WordNet é mencionada no texto, faz-se referência a WN de Princeton, que deu origem às demais wordnets desenvolvidas pelo mundo.

Com o desenvolver das pesquisas em torno da WN, psicólogos e linguistas de Princeton, ainda na década de oitenta, empenharam-se no desenvolvimento de uma ferramenta que organizasse o léxico de acordo com a relação estabelecida entre as palavras. Nesse recurso, o léxico se divide nas quatro principais classes de palavras lexicais: nome, verbo, adjetivo e advérbio. Inicialmente, a WordNet utilizava apenas nomes; depois, verbos e adjetivos foram adicionados; e, por último, incluíram-se advérbios. Todos esses dados se organizam através de um conjunto de sinônimos, os *synsets* (*synonym sets*). Cada *synset* representa um conceito lexical subjacente, com a pretensão de alcançar todas as lexicalizações de um conceito, recuperando, em tese, um mesmo sentido básico, inicial.

Diferentemente da FrameNet, nesse banco de dados não há uso de *corpus* para a constituição de *synsets*. Contudo, Fellbaum e Miller (2007) consideram que a definição disponibilizada para cada um bem como os exemplos resultam em um *corpus* semanticamente anotado, útil para sistemas de testes de aprendizagem automática. Acerca das definições, ou glosas, a FrameNet e a WordNet apresentam propósitos análogos.²⁶ Vossen (2002) afirma que, no sistema das wordnets, os conceitos se definem na relação com outros conceitos. Desse modo, as glosas, como Marrafa (2001) avalia, não apresentam o intuito de atuar como definições propriamente ditas, mas, sim, disponibilizar informações complementares, sem o interesse de determinar o sentido como fazem os dicionários tradicionais. Em alguma medida, a FrameNet tem concepção semelhante: ainda que exista o claro interesse em definir os *frames*, e que as Unidades Lexicais disponham de uma definição no banco de dados, bem como os Elementos de *Frame*, o objetivo desse recurso não é estabelecer uma definição para o usuário, mas permitir que, a partir das relações entre ULs, *frames* e EFs, o usuário seja capaz de estabelecer o sentido apropriado.

Levando em consideração o uso do dicionário para não especialistas, a WN mantém uma interface que guarda mais traços com os dicionários eletrônicos tradicionais que a FN – Figura 29. A primeira característica percebida é a divisão entre as classes de palavras, no caso, nominal e verbal. Dentro de cada uma, há disponível, em hiperlink, a estruturação das relações semânticas do *synset*. Ao lado

²⁶ Ainda que tanto a FrameNet quanto a WordNet possam ser comparadas a um *thesaurus*, já que ambas organizam o léxico em torno da similaridade semântica, a FN os reúne em torno de *frames*, e a WN o faz através de *synsets*.

da palavra procurada, há outras consideradas sinônimas a fim de que o usuário possa transitar pelas informações relacionadas. Nesse processo, as glosas atuam como facilitadoras, pois indicam o contexto em que cada *synset* vai se formar. Dentre os recursos disponíveis, em **Display Options**, Figura 29, há como o usuário escolher quais dados ele quer obter de retorno, por exemplo, com glosas ou sentenças contextualizadas.

The screenshot shows the WordNet Search interface. At the top, it says "WordNet Search - 3.1" with links to "WordNet home page", "Glossary", and "Help". Below this is a search bar with the word "visit" entered and a "Search WordNet" button. There are also "Display Options" with a dropdown menu and a "Change" button. A key explains that "S:" stands for "Show Synset (semantic) relations" and "W:" for "Show Word (lexical) relations". Below this, it says "Display options for sense: (gloss) 'an example sentence'".

Noun

- **S: (n) visit** (the act of going to see some person or place or thing for a short time) *"he dropped by for a visit"*
- **S: (n) visit** (a meeting arranged by the visitor to see someone (such as a doctor or lawyer) for treatment or advice) *"he scheduled a visit to the dentist"*
- **S: (n) visit** (the act of visiting in an official capacity (as for an inspection))
- **S: (n) visit** (the act of going to see some person in a professional capacity) *"a visit to the dentist"*
- **S: (n) sojourn, visit** (a temporary stay (e.g., as a guest))

Verb

- **S: (v) visit, see** (go to see a place, as for entertainment) *"We went to see the Eiffel Tower in the morning"*
- **S: (v) travel to, visit** (go to certain places as for sightseeing) *"Did you ever visit Paris?"*
- **S: (v) visit, call in, call** (pay a brief visit) *"The mayor likes to call on some of the prominent citizens"*
- **S: (v) visit, inspect** (come to see in an official or professional capacity) *"The governor visited the prison"; "The grant administrator visited the laboratory"*
- **S: (v) inflict, bring down, visit, impose** (impose something unpleasant) *"The principal visited his rage on the students"*
- **S: (v) chew the fat, shoot the breeze, chat, confabulate, confab, chitchat, chit-chat, chatter, chaffer, natter, gossip, jaw, claver, visit** (talk socially without exchanging too much information) *"the men were sitting in the cafe and shooting the breeze"*
- **S: (v) visit** (stay with as a guest) *"Every summer, we visited our relatives in the country for a month"*
- **S: (v) visit** (assail) *"He was visited with a terrible illness that killed him quickly"*

Figura 29: Busca pelo lema *visit* na *WordNet*

Fonte: <<http://wordnetweb.princeton.edu>>

O propósito de agrupar as palavras em *synsets* contribui para a análise de que essa rede se constrói a partir de uma abordagem paradigmática para a descrição lexical, uma vez que explora conjuntos de palavras capazes de produzir sentidos similares em dado contexto. Assim sendo, na sentença (33), *visitar* poderia ser trocado por *conhecer*, visto que esses verbos estruturam um mesmo *synset* através da relação de sinonímia; e, em (33), os pares se constroem a partir da relação de hiponímia / hiperonímia, uma vez que *beleza natural* é um tipo de *atração turística*.

- (33) Deveria ser proibido ir a Foz do Iguaçu e não visitar o Parque das Aves. Um exemplo de preservação, cuidado e carinho com espécies ameaçadas de extinção e um deslumbre aos olhos. (Andarilhos_do_Mundo)
- (34) Conhecida como “Cidade Maravilhosa”, as belezas naturais e a hospitalidade da população do Rio de Janeiro merecem destaque. (BrasilGov)

Diante de exemplos como esses, Miller et al. (1993) adotam uma noção mais fraca de sinonímia, aquela relativa ao contexto, a “similaridade semântica”. Com essa proposição, os autores reconhecem que não devem ser buscados sinônimos perfeitos, entretanto ressaltam que se tratam de relações simétricas, “se X é semanticamente similar a Y, então Y é igualmente similar a X”. Os motivos que levam a WordNet a submeter a organização dos *synsets* a categorias de palavras passa, necessariamente, pelo conceito da sinonímia. Como os *synsets* são a representação de um conceito através de um conjunto de palavras sinônimas, os autores avaliam que não se pode, por exemplo, trocar um substantivo por um adjetivo sem que a compreensão seja modificada. Por esse motivo, a WordNet organiza os dados a partir das classes de palavras lexicais, como é observado na Figura 29. Essa divisão contribui, segundo os autores, para a hipótese psicolinguística de que nomes, verbos e modificadores se dispõem de modo independente na memória semântica, resultando em redes autônomas (MILLER et al., 1993, p.7). Diante da FrameNet, esse é um ponto importante para distanciá-las, já que, para a FN, o intuito é agregar todas as palavras, sejam elas verbos, nomes, adjetivos, advérbios ou preposições, que evoquem um *frame*.

Além da sinonímia, que, certamente, é a relação mais importante da WordNet, base para a constituição de *synsets*, a sistematização de relações estabelecidas por relações como É TIPO DE e É PARTE DE são conceitos semânticos básicos para que as análises aconteçam. A antonímia é uma relação trabalhada pela WN, que, embora tenha uma estrutura similar à sinonímia, apresenta suas particularidades. Segundo Miller et al. (1993), a indicação psicolinguística mais eficaz para descobrir pares de palavras antônimas se dá através de testes que utilizem a intuição das pessoas: pergunta-se, por exemplo, qual a primeira palavra que vem à mente quando se ouve *victory* (*vitória*), a maioria das pessoas responde *defeat* (*derrota*), e *victory* quando questionadas acerca de *defeat*. Entretanto, os autores ponderam que essa tarefa não é assim tão simples quanto possa parecer. Embora eles estabeleçam que a antonímia seja uma relação semântica entre a forma das

palavras, incorpora-se, na estruturação dessa relação, uma oposição conceitual entre os pares de palavras. Assim os *synsets* {*rise; ascend*} (*subir; ascender*) e {*fall; descend*} (*cair; descer*) estabelecem uma oposição conceitual, mas não devem ser considerados pares de antônimos diretos.

(35) Milhares de brasileiros ascenderam socialmente com o governo Lula.

O verbo *ascender*, no exemplo ilustrado em (35), tem como antônimo *descender*, e, embora nesse contexto, *cair* seja uma opção aceita, esse verbo não é considerado antônimo de *ascender*, mas um oposto conceitual. Para sustentar essa colocação, eles ressaltam que, quando as pessoas são questionadas acerca dos pares {*rise; ascend*} e {*fall; descend*} como antônimos, elas hesitam em concebê-los como tais. Isso, para os autores, enfatiza que existe uma oposição entre a forma dos pares *rise* e *fall* e *descend* e *ascend*.

Além dessa questão enfrentada na sistematização de *synsets* antônimos, outro desafio é trazido pela postulação de que o antônimo de *x* é não *x* (MILLER ET AL., 1993, p. 7). Há vários pares, considerados antônimos, que atestam casos desse tipo. Algo que seja não quente não, necessariamente, é frio; a palavra *morno* mostra a existência de algo que se situa entre os dois extremos, evidenciando que é possível graduar a relação. Outro exemplo é o par *solteiro* e *casado*, que, tradicionalmente, implicava uma condição excludente. Entretanto, hoje, institucionalmente, já se validam conceitos intermediários, como o estabelecimento da união estável, que legitima uma relação intermediária entre os dois *status*. Além desses exemplos que evidenciam a limitação do tratamento da antonímia pelos moldes da lógica aristotélica, o exemplo de *bachelor*, amplamente abordado pela literatura em Linguística Cognitiva (cf. LAKOFF, 1987), sugere que, embora um padre seja não casado, ele não pode ser considerado um *solteirão*, como sugere a palavra *bachelor*, pois um padre não está disponível para o matrimônio, impossibilitando, assim, sua participação em uma categoria que designa aqueles que chegaram a uma determinada fase da vida e ainda não se casaram.

O tratamento dado pela WordNet acerca dos adjetivos, além de estar inserido na relação de sinonímia, é especial na antonímia. Dada a diversidade dessa classe, a WN segue a divisão feita entre adjetivos relacionais e adjetivos descritivos. Em linhas gerais, para tratar dos adjetivos considerados relacionais, é retomado o conceito “relativo a / pertence a, associado com” atribuído a Levi (1978 apud MILLER ET AL., 1993). A partir dessa definição, o adjetivo *dental* advém do

substantivo *dente* – relativo a –, conforme definição. Os adjetivos considerados relacionais, geralmente, são derivados de formas nominais, não possuem antônimos e não são graduáveis; por isso, não se costuma ouvir graduações do tipo *creme muito dental*.

Já a classe descritiva os organiza de acordo com os atributos, que podem ser divididos por semelhança de significado (sinonímia) ou em termos de oposição binária (antonímia); quando não houver um antônimo considerado direto, ocorrerá um empréstimo entre aqueles que os têm. A Figura 30 mostra o exemplo de um conjunto de adjetivos que se localizam em torno de *molhado* e *seco*. Palavras como *aguado* e *ressecado* mantêm uma relação que Miller et al., (1993, p. 29) considerariam como “antonímia indireta”, já que não se pode fazer uma relação de oposição binária entre os pares.

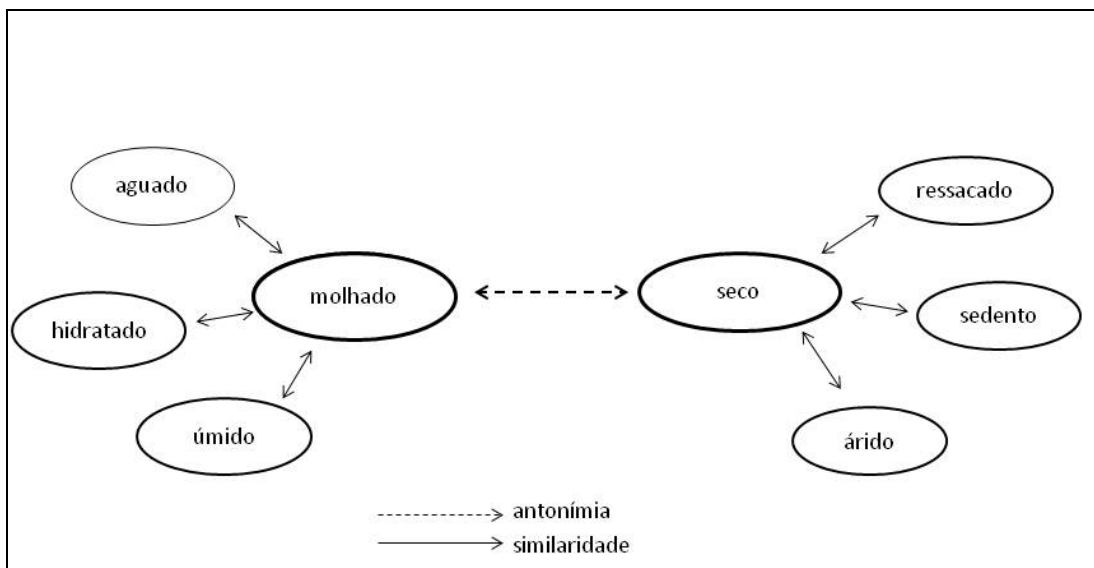


Figura 30: Estrutura de adjetivos bipolares

Fonte: Adaptação do exemplo extraído de Miller et al., 1993, p.29

Diante de toda a discussão por trás deste recorte, percebe-se a dificuldade de fazer generalizações a respeito da noção de simetria na antonímia. Cruse (1986) aborda esse tópico mostrando que, na língua inglesa, poucos são os pares que guardam uma simetria formal, como, por exemplo, *increase / decrease*, *accelerate / decelerate*. Como expõe a Figura 30, afirmar que a antonímia seja uma relação lexical estabelecida estritamente entre a forma das palavras não é tão simples.

Na constituição do banco de dados da WordNet, além da sinonímia e da antonímia, outras relações que merecem destaque são hiponímia / hiperonímia e

meronímia / holonímia. Estas estruturam um grupo de palavras que se agrupam em torno de seus significados, sendo então consideradas relações semânticas, diferentemente do que ocorre com a antonímia, como Miller et al. (1993) apontam. Pela relação hiponímia / hiperonímia, em certa medida²⁷, é possível sistematizar uma hierarquia entre os pares, já que é marcada pela noção de subordinação e superordenação. Vejam-se os exemplos (36) e (37).

- (36) Vale lembrar que os estudos do Ministério do Turismo apontam para um número de 600 mil turistas estrangeiros durante a Copa do Mundo da FIFA em 2014. (FIFA)
- (37) O melhor dessa pousada é que fica do lado do terminal de ônibus que te leva pra fazer todos os passeios possíveis de Iguaçu. Saem ônibus para o Paraguai, Cataratas Brasileiras, Argentinas, para Itaipu, pra mochileiro é ótimo. No primeiro dia peguei o busu para Cataratas Brasileiras (...) (Destino_de_viagem).

Analisando os exemplos em (36) e (37), pode-se dizer que *mochileiro* carrega todas as propriedades básicas de *turista*, mas se distingue ao especificar traços de um grupo particular, daqueles que, geralmente, viajam com pouca bagagem e procuram meios mais econômicos de transporte e hospedagem, por exemplo. Assim, *mochileiro* é hipônimo de *turista*, e *turista* é hiperônimo de *mochileiro*. Ainda que se possa afirmar que um mochileiro seja um turista, o contrário não procede. Dessa forma, não há espaço para dizer que se trata de uma relação simétrica. Desconsiderando-se a diferença de escopo entre os bancos de dados, a relação hiponímia / hiperonímia guarda traços com a relação Herança estruturada pela FrameNet, uma vez que, de modo geral, estabelecem-se pela relação TIPO DE.

Outra relação é a meronímia / holonímia, conhecida também como relação parte / todo. Um *synset* que representa essa relação é o par *pneu* e *motor*, que são merônimos (parte) que se relacionam com *carro* holônimo (todo). Diz-se então que A (*pneu; motor*) é parte de B (*carro*), e B (*carro*) inclui A (*pneu; motor*). A mesma analogia pode ser realizada com *braço / corpo*, *dedo / mão*. Contudo, conforme pondera Marrafa (2002), essa relação abarca uma “complexa família de relações”. Um exemplo é o par *flor* e *planta*, pois ainda que flor seja parte de uma planta, esta não, necessariamente, inclui uma flor.

Tendo em vista que o léxico seja, como Fontenelle (2000 apud SCHRYVER, 2003. p.188) avalia, um repositório de milhares de conceitos e palavras ligadas em torno de uma imensa rede, a WordNet, por meio dos *synsets*, opta por organizá-la

²⁷ Marrafa (2002) aprofunda a discussão da relação hiponímia / hiperonímia, inserindo casos que dificultam analisar os níveis de hierarquia entre os pares.

através das relações lexicais. Como foram assinalados nesta seção, diversos casos são contemplados, mas vários obstáculos trazidos pela metodologia adotada são enfrentados, uma vez que a relação estabelecida entre as palavras é mais complexa do que se pode sistematizar com tais relações.

2.3.2 Expansão da WordNet para Outras Línguas

Assim como a FrameNet, a WordNet é eficaz não só como dicionário para atender a usuários humanos mas também como banco de dados para diversas aplicações computacionais. Fellbaum e Miller (2007) ressaltam que, passadas poucas décadas de sua criação, a WordNet é hoje amplamente utilizada em PLN. Como os autores observam, o uso do comando DEFINE no Google retorna glosas e frases de exemplo da WN, refletindo sua popularidade.

Atualmente, existem wordnets em mais de quarenta línguas, incluindo aquelas como latim e sânscrito. Para o português do Brasil, há a WordNet.Br²⁸, uma iniciativa do pesquisador Bento Carlos Dias da Silva, membro do corpo docente da Universidade do Estado de São Paulo, campus Araraquara. Essa rede lexical em andamento já conta com a base de dados verbais, totalizando cerca de 5.900 verbos em mais de 3.700 *synsets*, que se alinham aos dados da WordNet de Princeton. Esse alinhamento é possibilitado através do índice interlingual da WordNet mãe.

Nesse sentido, a EuroWordNet (VOSSSEN, 2002) também se estrutura como uma iniciativa multilíngue. Criada na década de noventa para representar oito línguas da União Europeia, ela expõe semelhanças e especificidades em relação à WN de Princeton. Como o intuito é sistematizar várias línguas, isso implica registrar diferentes realidades. Por esse motivo, são somadas ao banco de dados relações semânticas e lexicais não discutidas na plataforma mãe. Como Vossen (2002) ressalta, o intuito é manter as diferenças culturais e linguísticas nas wordnets.

Para que seja possível esse empreendimento, a ligação entre essas wordnets acontece através do chamado *Inter-Lingual-Index* (ILI); o intuito é que todos os conjuntos de *synsets* nessas línguas apresentem um equivalente no ILI para que a correspondência aconteça.

²⁸ Para mais informações sobre a WordNet.Br: <http://caravelas.icmc.usp.br/wordnetbr/>

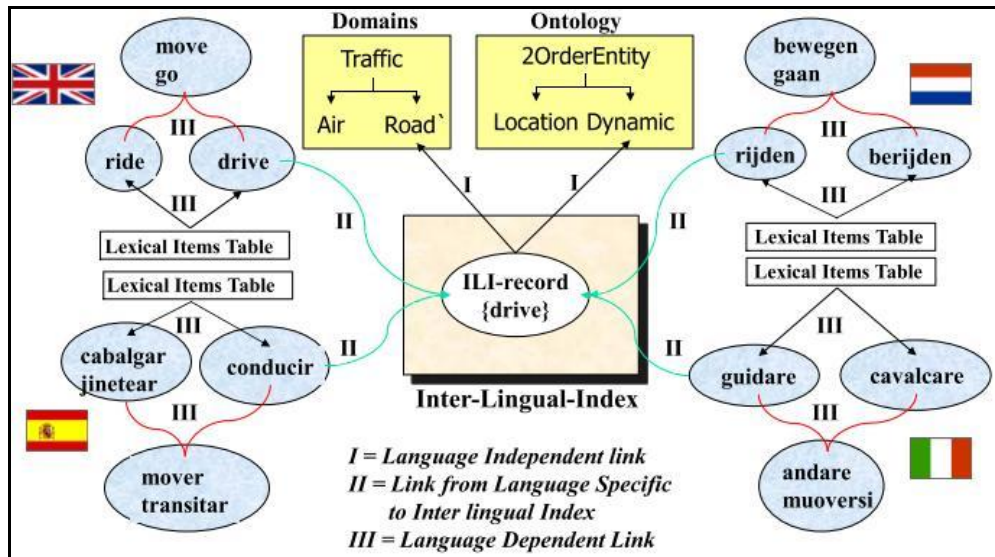


Figura 31: Inter-lingual-Index na base de dados da EuroWordNet
 Fonte: <<http://www.vossen.info/>>

A Figura 31 representa como se dá essa interlíngua. O exemplo é construído com *synsets* verbais da ação de conduzir em quatro línguas diferentes. Observa-se que cada *synset* possui um representante (*rijden*, *guidare*, *conducir*, *drive*) que se associa ao registro no ILI, no caso, *drive*. O objetivo final é a representação de várias lexicalizações distintas para um mesmo conceito. Além de ser um rico recurso lexical para usuários e pessoas interessadas em processamento de linguagem natural, de modo semelhante à FrameNet, esse recurso permite que especialistas em linguagem possam estudar os padrões de lexicalização das línguas.

Além dessa iniciativa, aplicativos interessantes vêm sendo desenvolvidos em torno da WordNet. Na língua portuguesa, há, por exemplo, o Lextec (MARRAFA ET AL., 2009) e o TemaNet (MARRAFA ET AL., 2006),²⁹ coordenados pela professora Palmira Marrafa, na Universidade de Lisboa. O Lextec é um projeto que visa a construir léxicos técnicos organizados através de *synsets*. Atualmente, contém dez domínios, dentre eles o Turismo, Figura 32.

²⁹ É importante agradecer à professora Palmira Marrafa por ter apresentado tais projetos e sugerido leituras para a realização desta seção.

O Lextec-Léxico Técnico do Português está a ser desenvolvido no âmbito de um projecto em curso - cujos resultados vão sendo disponibilizados progressivamente - e inclui os seguintes recursos:

- um glossário que integra expressões linguísticas de domínios específicos, bem como as respectivas definições;
- uma base de textos autênticos que ilustram os contextos de uso das expressões em questão e/ou fornecem informações adicionais de interesse para a compreensão dos conceitos envolvidos;
- uma rede conceptual que relaciona entre si as expressões com o mesmo significado, que, por seu turno, são relacionadas com outras, como acontece, por exemplo, com as de significado mais geral ou as de significado mais específico; a rede especifica ainda, para todas as expressões, relações de correspondência com expressões em Inglês.

Para referir este trabalho:

Marrafa, P., R. Amaro, S. Mendes, R. P. Chaves & S. Lourosa (2009) *Lex Tec – Léxico Técnico do Português: Ambiente, Banca, Comércio, Construção Civil, Direito Comercial Internacional, Economia e Gestão de Empresas, Energia, Seguros, Turismo, Telecomunicações*. Instituto Camões, <http://instituto-camoes.pt/lextec>.

Figura 32: Lextec - Léxico Técnico do Português

Fonte: <<http://instituto-camoes.pt/lextec>>.

Cada área contemplada apresenta um glossário com cerca de mil palavras ou expressões consideradas produtivas em cada domínio; textos que ilustram essas expressões em uso; e/ou informações adicionais de tais conceitos, incluindo correspondências para o inglês. A rede lexical estabelecida pelas wordnets é responsável por consolidar as principais relações entre os significados nesse banco de dados. Já o TemaNet é um recurso de domínio semântico específico, cujo principal objetivo é auxiliar nos contextos de ensino / aprendizagem de língua portuguesa. Abrange doze domínios, dentre eles o esportivo.

bola (cat. nominal)

glosa: artigo desportivo, tipicamente esférico, usado para obter pontos em determinados desportos

exemplo: Muitos desportos são jogados com bolas específicas.

equivalência em Inglês

<p>"bola" é subtipo de</p> <ul style="list-style-type: none"> artigo desportivo 	<p>são subtipos de "bola"</p> <ul style="list-style-type: none"> bola de futebol, esférico bola de rãguebi, bola de rugby bola de andebol bola de basquetebol, bola de basquete, bola de basket bola de golfe, bola de golf bola de voleibol, bola de vôlei, bola de volley bola de ténis bola de ping-pong, bola de pingue-pongue, bola de ténis de mesa bola de bilhar bola de bowling bola medicinal bola de snooker bola de hóquei bola de pólo aquático bola de squash bola de críquete bola de futebol americano bola de softbol, bola de softball bola de baseball, bola de beisebol 	<p>relações não-hierárquicas para "bola":</p> <ul style="list-style-type: none"> tem como característica ser <ul style="list-style-type: none"> esférico (mas não necessariamente) co_relaciona-se com <ul style="list-style-type: none"> raquete (mas não necessariamente) está envolvido em <ul style="list-style-type: none"> finta (mas não necessariamente) lance de bola parada (mas não necessariamente) golo (mas não necessariamente) corfebol (mas não necessariamente) pólo (mas não necessariamente) está implicado como objecto em <ul style="list-style-type: none"> serviço (mas não necessariamente) passar (mas não necessariamente) passar (mas não necessariamente) passar (mas não necessariamente)
--	--	---

Figura 33: Resultado da busca pelo lema *bola* no TemaNet

Fonte: <<http://www.instituto-camoes.pt/temanet/>>.

A Figura 33 mostra o resultado da busca pelo substantivo *bola* dentro do domínio esporte. Foram retornadas as seguintes informações: a categoria gramatical, no caso, nominal; uma glosa, indicando que se trata de um objeto tipicamente esférico, utilizado para obter pontos em determinados esportes; um possível exemplo da palavra em uso, com um link do termo equivalente na língua inglesa; por último, as principais relações lexicais envolvendo esses lexemas. Dentre elas, as primeiras foram aquelas estruturadas através da hierarquia TIPO DE. Dessa forma, mostrou-se que *bola* é um subtipo de *artigo esportivo* (*bola_hipônimo / artigo_esportivo_hiperônimo*), que carrega distintas características de acordo com os tipos de esporte (*bola_hiperônimo / bola_de_futebol_hipônimo*). Também foram consideradas relações não hierárquicas, como a característica de ser esférica e a relação que mantém com raquete.

2.4 O Kicktionary: Um Dicionário Multilíngue para O Futebol

Recurso lexical multilíngue da linguagem do futebol, o Kicktionary é um dicionário temático, disponível *online* gratuitamente em <<http://www.kicktionary.de/>>, que abrange as línguas inglesa, alemã e francesa. O objetivo principal era construir um dicionário que se destacasse dos demais. Para isso, além da tecnologia de hipermídia, procurou-se explorar como teorias linguísticas sobre semântica lexical, somadas aos métodos da Linguística de *Corpus*, poderiam contribuir para a construção desse produto (SCHMIDT, 2006, 2007, 2008, 2009).

O *corpus* utilizado pelo Kicktionary para ilustrar os exemplos e sustentar a análise é, em grande parte, paralelo e advém das partidas de futebol coletadas do site oficial da UEFA, associação europeia de futebol. Para o alemão, há material adicional, coletado do jornal Kicker, além de partidas narradas no rádio. Atualmente, o dicionário contém cerca de 2.000 Unidades Lexicais, compreendidas entre substantivos, verbos, adjetivos e expressões idiomáticas. Para cada sentença, há até dez frases anotadas. Várias se constroem em torno de verbos suporte, como *fazer substituição* [*make substitution*] e *cometer falta* [*commit foul*].

A utilização da Semântica de *Frames* (FILLMORE, 1982, 1985) bem como a metodologia da FrameNet (FILLMORE ET AL., 2003, 2003a; RUPPENHOFER ET AL., 2010) tiveram destaque neste processo, assegurando ao Kicktionary o prestígio

de ser a primeira tentativa concreta de dicionário temático organizado em *frames*. Nessa elaboração, a WordNet (MILLER ET AL., 1993; FELLBAUM, 1998) também apresenta seu mérito, uma vez que o tratamento das relações lexicais do domínio do futebol utilizam um tipo especial de *synset*. Em decorrência da utilização de conceitos básicos que permeiam a constituição de ambas as ferramentas, um tratamento mais refinado do comportamento semântico-lexical do vocabulário do futebol é possibilitado. São esses conceitos a cena, o *frame*, a Unidade Lexical e o *synset*.

Embora seja abordado por Fillmore (1977), o conceito de cena, que é amplamente explorado no Kicktionary, não é empregado na FrameNet. A cena, conforme Fillmore explicitava neste texto, podia ser entendida como uma estrutura do conhecimento e da experiência do ser humano que, diferente do *frame*, não precisa ser manifestada linguisticamente. Por isso, Schmidt (2009) vai realçar que o *frame* é uma entidade estrutural utilizada para agrupar Unidades Lexicais que dividem o mesmo significado básico em dado evento, demarcando uma perspectiva similar e possibilitando que as relações semânticas dos seus argumentos sejam generalizadas. Já a cena corresponde, no Kicktionary, a eventos prototípicos das partidas de futebol, que reúnem diversos *frames*. Nesse sentido, a cena é a construção superordenada de um *frame*, responsável por lidar com as propriedades dos significados linguísticos concretos de tal domínio.

On the pitch	
Shot 12 frames 269 lexical units	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) Save (46 LUs) Shoot_At (5 LUs) Shot (123 LUs) Shot_Supports (25 LUs)
Pass 11 frames 226 lexical units	Bad_Pass (2 LUs) Being_Free (14 LUs) Connect (5 LUs) Control (37 LUs) Flick_On (7 LUs) Intercept (11 LUs) Mark (20 LUs) Pass_Back (5 LUs) Pass_Combination (21 LUs) Supply_Pass (12 LUs)
Goal 10 frames 148 lexical units	Award_Goal (11 LUs) Celebrate_Goal (5 LUs) Concede_Goal (7 LUs) Convert_Chance (8 LUs) Goal (85 LUs) Multiple_Goals (1 LUs) Overcome_Goalkeeper (7 LUs) Own_Goal (4 LUs) Prepare_Goal (6 LUs) Score_Goal (7 LUs)
One_On_One 7 frames 129 lexical units	Beat (49 LUs) Challenge (18 LUs) Deny (19 LUs) Lose_Ball (7 LUs) One_On_One (25 LUs) Take_On (5 LUs) Trick (6 LUs)
Foul 12 frames 211 lexical units	Advantage (4 LUs) Concede_Compensation (4 LUs) Dissent (12 LUs) Foul (54 LUs) Give_Card (8 LUs) Offside (8 LUs) Referee_Decision (12 LUs) Sanction (36 LUs) Set_Piece (45 LUs) Simulation (6 LUs) Win_Compensation (6 LUs)

Figura 34: Visualização de *On the Pitch* no Kicktionary

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

Nesse dicionário, há a descrição de 16 cenas, compreendidas dentro de três superdomínios: *On the Pitch* (No Gramado), *A match in a competition* (Partida em

Competição) e *Actors and Objects* (Atores e Objetos). Na Figura 34, o superdomínio *No Gramado* é ilustrado com algumas de suas cenas. A *One_on_One* (Um_a_um) acontece quando dois jogadores competem entre si pela posse da bola. Há sete *frames* descritos para essa situação, dentre eles, *One_on_One* (Um_a_um), *Challenge* (Desafio) e *Take_On* (Posse_da_bola). Quando a perspectiva está centrada na disputa em si, tem-se o *frame Um_a_um*, se se tem em foco aquele que disputa a bola em posse de outro jogador, essa perspectiva é descrita em Desafio, mas, se o objetivo é analisar o jogo a partir do ponto de vista daquele que tem a posse da bola, o *frame* é *Posse_da_bola*.

Os exemplos em (38), (39) e (40) descritos no Kicktionary mostram como esses *frames* se comportam: em (38), embora os jogadores sejam argumentos centrais, o foco não está neles, mas nos acontecimentos que permeiam a disputa; em (39), a análise acontece a partir do jogador que tenta ter a posse da bola; e, por último, em (40), a perspectiva do jogador com a bola é o foco para a análise da disputa. Para cada um dos *frames*, há várias Unidades Lexicais que o evocam.

- (38) [Markus Babbel da Estugarda JOGADOR_1] disputa [a bola no ar BOLA_EM_MOVIMENTO] [com Daniel Hested JOGADOR_2].
- (39) [Jorge Costa do Porto JOGADOR_OPONENTE] desafia [Ivica Olic JOGADOR_COM_A_BOLA].
- (40) [Thomas Gravesen da Dinamarca JOGADOR_COM_A_BOLA] assume [a defesa ucraniana JOGADOR_OPONENTE]

No *frame Challenge* (Desafio), a Figura 35 destaca a existência de 18 ULs, somando-se as do alemão, inglês e francês.

Lexical Units / Lexikalische Einheiten / Unités lexicales	
🇩🇪	abdrängen angreifen Attacke attackieren bedrängen Einsteigen Grätsche in_die_Zange_nehmen stören Tackling
🇺🇸	challenge challenge sliding_tackle tackle tackle
🇫🇷	tacle glissé tacle tacler

Figura 35: Unidade Lexical do *frame Challenge*
 Fonte: <<http://www.kicktionary.de>>.

Além dos conceitos de cena, *frame* e UL, o de *synset* também deve ser mencionado. No dicionário, as ULs são organizadas em *synsets*, que são também responsáveis por estabelecer as relações hierárquicas encontradas no vocabulário do futebol: hiponímia / hiperonímia, holonímia / meronímia e troponímia. Entretanto, Schmidt expande essa função, previamente prevista pela WN. Para ele, a noção de *synset* irá agrupar não apenas ULs sinônimas mas também os possíveis

correspondentes que o analista irá indicar nas três línguas. Assim, no *frame* *Celebrate_Goal* (*Celebrar_Gol*), as ULs *bejubeln*, *feiern* (alemão); *célébrer*, *fêter* (francês); e *celebrate* (inglês) pertencem a um mesmo *synset*. Ao ampliar o conceito de *synset*, a opção de Schmidt foi realizar a tradução através dos recursos da WordNet, diferentemente do Copa 2014, que opta por realizar as traduções via *frames*, como é apresentado no capítulo 4.

No Brasil, há o empenho na construção da contraparte do Kicktionary para o português do Brasil, o Kicktionary_Br, organizado pela professora Rove Chishman na Universidade do Vale dos Sinos, em São Leopoldo.

3 O CENÁRIO DO TURISMO

Eventos como a Copa do Mundo FIFA 2014 proporcionam um tipo específico de turismo, conhecido como turismo de evento. A inclusão do domínio do turismo no dicionário Copa 2014 foi influenciada pela relevância desse ramo de atividade no Brasil. O fato de os 8.547.403 km² ocupados pelo país incluírem um panorama cultural e natural bastante diversificado, com paisagens belíssimas e diferentes costumes, faz com que os milhares de turistas tanto estrangeiros quanto brasileiros que estarão nas cidades-sede para assistirem aos jogos da Copa do Mundo FIFA 2014 optem por conhecer tais atrações.

A iniciativa do Copa 2014 FrameNet Brasil tem como intuito auxiliar turistas e também aqueles que estarão envolvidos na organização do evento. Seja como visitantes ou guias, milhares de pessoas irão se comunicar, em especial no que tange ao potencial turístico não só das cidades que sediarão o evento como também dos arredores. Dessa forma, a tentativa é sistematizar, com o suporte teórico trazido pela Semântica *Frames* e com apoio de *corpora* especializados, o cenário amplo e diversificado que estrutura o domínio turístico. Considerando que a inserção desse domínio para o dicionário deva ser vista como um guia de usos linguísticos, pretende-se que, a partir dos exemplos das entradas lexicais, o usuário seja direcionado para os sites que forneceram tais dados. Assim, a busca específica do usuário o direcionará a diversos repositórios de informações de fomento a atividade turística como sites dos órgãos governamentais brasileiros e blogs de viajantes.

Diante de tais objetivos, este capítulo se organiza da seguinte maneira: a seção 3.1 insere discussões teóricas do Turismo bem como a estreita relação que essa área estabelece com outros domínios; a seção 3.2 apresenta o processo de criação do cenário do turismo nos moldes da FrameNet Brasil; a seção 3.3 trata da descrição dos *frames* de experiência turística analisados e Unidades Lexicais que evocam tais *frames*.

3.1 Turismo como Atividade Humana

Especificar ao certo onde e quando ocorreram as primeiras atividades turísticas no mundo não é uma tarefa fácil. Vários autores apresentam dados

distintos, que demonstram a dificuldade de alcançar um consenso nesse sentido. O esforço de Silva e Kemp (2008) para demonstrar a evolução histórica dessa atividade sinaliza que, desde as civilizações antigas, milênios antes de Cristo, já se viam traços de atividades turísticas. Por mais que as metas das viagens não tivessem como foco explorar novas culturas e cenários geográficos, esse resultado era consequência de tais deslocamentos, o que fazia com que os viajantes retornassem aos seus locais de origem compartilhando várias experiências vivenciadas.

A etimologia da palavra *turismo*, como diz Moesch (2002 apud TOSQUI, 2007), remonta ao latim com o substantivo *tornus* (*torno*) e o verbo *tornare* (*girar, redondear*), que dão ideia de uma viagem que volta ao local de partida. Já a raiz *tour* vai aparecer, conforme os dados desse autor, na Inglaterra no século XVIII. Entretanto, conforme Trigo (1998 apud SILVA & KEMP, 2008) avalia, será apenas em meados do século XIX que o turismo surge como prática organizada, já que o desenvolvimento tecnológico trazido pela Revolução Industrial bem como o surgimento da burguesia possibilitaram a difusão dessa prática na sociedade.

Um desafio ainda maior nesse campo de estudo é o acordo em torno de um conceito que defina a palavra *turismo* e, conseqüentemente, a tarefa de delimitar quem seria o *turista*. Uma breve pesquisa por tais definições mostrou que a tentativa de estabelecer o escopo dessa atividade bem como o perfil de seu praticante se confronta com inúmeras limitações (cf. COOPER ET AL., 2002; TOSQUI, 2007). Afirmar que essa viagem é alheia a interesses profissionais ou ainda a necessidade de o turista ser proveniente de país estrangeiro são assertivas que negligenciam vários dados da realidade.

Para Cooper et al. (2002, p. 36), vários mitos, como o que considera o turismo apenas como atividades de lazer realizadas durante as férias, garantiram a imagem do *glamour* ao qual as pessoas ainda associam essa atividade. A dificuldade de estabelecer um consenso na definição desses conceitos reflete, para esses autores, não só a complexidade da área mas indica também como esse campo de estudo ainda é prematuro. Contudo, eles ponderam que a necessidade de formalizar determinadas terminologias, além de ser uma tentativa para garantir a credibilidade da área, é importante por questões práticas de medição e legislação (COOPER ET AL., 2002, p. 42).

Ainda que esse objetivo siga em andamento, cada definição vai deixando algum aspecto interessante, destacando traços que contribuem para delinear o perfil do turista e o escopo da atividade. Por exemplo, enfatizar a necessidade do deslocamento fora do ambiente de residência; especificar que, no turismo, não há o objetivo em fixar residência e, por isso, o fato de que costuma ser temporário; e ligá-lo a trocas de experiências culturais são todas características presentes nesse campo de estudo.

Uma proposta que dialoga com a temática deste trabalho é a tese de doutoramento de Patrícia Tosqui (TOSQUI, 2007). A autora constrói uma ontologia de domínio para contribuir na elaboração de um vocabulário bilíngue para o turismo com fins didáticos.

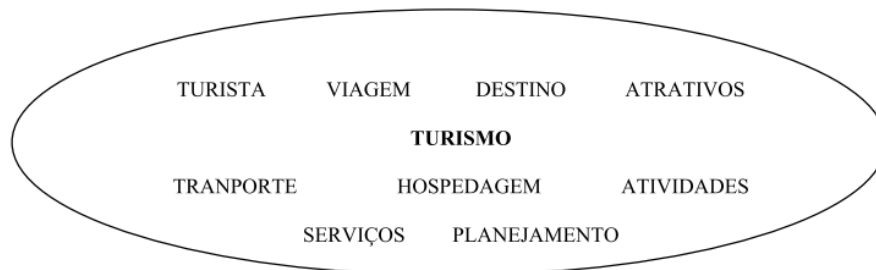


Figura 36: Caracterização informal dos conceitos do turismo
Fonte: TOSQUI, 2007, p. 88

O esboço inicial constituído por Tosqui para guiar a modelagem da ontologia, Figura 36, reflete a diversidade de domínios que estabelecem contato com o Turismo. Devido ao laborioso estudo contrastivo realizado para verificar como especialistas definiam essa área, a autora formula um conceito que, em suas palavras, foi “resultado de uma interpretação de fatos naturais e/ou culturais”.

[Turismo é] uma atividade humana realizada voluntariamente, quer individual ou coletivamente, que inclui uma viagem, com duração determinada, para fora do entorno habitual do turista, com fins de lazer, de negócios, médicos, culturais, entre outros, e que envolve produtos, serviços e atividades econômicas ligadas direta ou indiretamente a ela. Para que essa atividade aconteça, é preciso que haja planejamento, que o local visitado tenha condições estruturais para receber o turista e que a comunidade local esteja preparada para oferecer os serviços e produtos requisitados de forma sustentável dos pontos de vista econômico, ambiental, cultural. (TOSQUI, 2005, p. 86)

Dentre a pluralidade de definições encontradas na literatura, que Tosqui (2007) defende como sendo resultado do caráter multidisciplinar da área, nenhuma descartou que se trata de uma atividade humana que vem movimentando de forma significativa a economia de muitos lugares. Nesse sentido, as definições propostas

para os *frames* de experiência turística refletem tanto as discussões apresentadas como também são resultado do estudo realizado diretamente nos dados durante o processo de constituição de *corpus*.

É importante destacar que não se tem a pretensão de as definições trazidas por este trabalho solucionarem todos os desafios apresentados, pois esse estudo requer aprofundamento das variadas áreas interligadas e rigor capaz de responder questões de legislação e medição para várias finalidades. Entretanto, a tentativa de modelar em *frames* o domínio turístico, sistematizando as etapas desse evento, permite aos estudiosos da área analisar essa atividade de outras maneiras.

3.2 Criação do Cenário do Turismo para a FrameNet

Além das várias características que legitimam o turismo como uma área específica, que deve ser analisada com rigor científico, existe um contato intrínseco com outras áreas, como Artes, Arquitetura, Economia e Geografia. A separação que, por vezes, é necessária acontece para fins de formalização e ensino / aprendizagem, pois, como reconhecem Cooper et al. (2002, p. 40), todos os elementos desse sistema se mantêm interligados. Pela reflexão dos autores, “na verdade [esses elementos] estão entrelaçados, e o entendimento desse inter-relacionamento conduz à verdadeira compreensão do tema”.

Para iniciar o estudo pelos *frames* deste domínio, foi feita uma pesquisa inicial no banco de dados fornecido pela FrameNet. Apesar de vários *frames* dialogarem com o evento turístico, como `Experiencer_focus` (Foco_no_experienciador), `Hospitality` (Hospitalidade), `Infraestructure` (Infraestrutura), `Travel` (Viajar), `Visiting` (Visitar), `Providing_Logding` (Fornecer_Hospedagem), dentre outros, o único que fornecia características exclusivas sobre a atividade turística foi `Touring` (Turismo).

Definition:

A **Tourist** experiences an **Attraction**, a location with a unique history or other societally recognized individual character, with the goal of seeing and learning about it. Typically, the **Attraction** has a source of information about the **Attraction**--a guide, pamphlets, or displays.

We TOURED the Grand Canyon.

FEs:

Core:

Attraction [Phen]	The societally-recognized unique location that the Tourist experiences.
Semantic Type: Location	
Tourist [per]	The individual that is seeking an experience at an Attraction.
Semantic Type: Sentient	

Figura 37: Definição do *Frame Touring* e seus elementos nucleares³⁰

Fonte: <<https://framenet2.icsi.berkeley.edu>>

Pela definição, Figura 37, um turista conhece uma atração, sendo o objetivo central da visita a obtenção de informações ou apenas a apreciação da atração em questão. Algumas Unidades Lexicais correspondentes para o português que evocam esse *frame*, conforme definido pela FrameNet de Berkeley, são *turista*, *visitar*, *apreciar*, *desfrutar* e *conhecer*.

- (41) [Turistas _{TURISTA}] VISITAM ^{ALVO} [o arquipélago de Fernando de Noronha _{ATRAÇÃO}] [para conhecer suas lindas praias _{FINALIDADE}].

Como indicado pela sentença (41), esse *frame* perfila o turista para a representação desse evento. Na etapa de construir o *corpus*, foram atestadas várias outras ocorrências que indicavam que este *frame* não conseguia demonstrar por completo a extensão desse evento. Havia sentenças enfatizando a atração turística; em outras, o foco estava no lugar que possuía atrações; alguns exemplos expressavam relações com outras áreas, entretanto, não estava claro se compunham *frames* com turismo, como questões de infraestrutura, compras, reservas, além da intrínseca relação que estabeleciam com o domínio da visita e da viagem, de modo geral.

Na tentativa de esboçar o que estaria compreendido na estruturação do turismo, passou-se então a pesquisar como a FrameNet sistematiza *frames* que perfilam vários elementos diferentes e mantêm relações com outras áreas.

³⁰ A descrição completa do *frame Touring*, com a definição de todos os seus elementos, Unidades Lexicais e relação com outros *frames* bem como as adaptações realizadas serão abordadas na seção 3.3.2 deste trabalho.

3.2.1 Frames de Cenário

A FrameNet marca a existência de *frames* produtivos e/ ou que apresentam uma estrutura complexa com a criação de tipos especiais de *frames*, chamados *cenários*. O objetivo é orientar a sistematização das relações entre *frames*, possibilitando uma estruturação minuciosa de como acontece essa interação.

O *frame* *Visiting_scenario* (Cenário_da_Visita) é um exemplo deles, Figura 38. Segundo sua definição, um Agente e uma Entidade passam a estar num mesmo local. Para alguma finalidade específica, o Agente permanece com a Entidade, mas, depois de algum tempo, retorna para o ponto de partida.

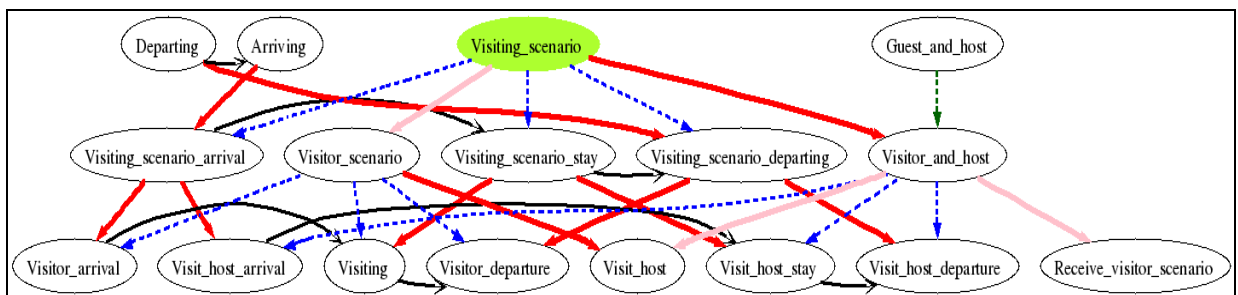


Figura 38: *Visiting_scenario*

Fonte: <<https://framenet.icsi.berkeley.edu/fndrupal/FrameGrapher>>

De acordo com essa estruturação, três subframes são construídos a partir do *frame* *Visiting_scenario*: *Visiting_scenario_arrival* (Cenário_da_visita_chegada), *Visiting_Scenario_stay* (Cenário_da_visita_estada) e *Visiting_scenario_departing* (Cenário_da_visita_partida). Tal subdivisão permite exibir que esse evento é marcado por etapas, conforme a sequência de acontecimentos: a chegada do visitante; sua estada junto à entidade visitada; e o retorno desse agente para o local inicial. Essa relação é conhecida como Precedência. Outra relação exibida por esse cenário é Perspectiva. Nessa, ainda que os *frames* compartilhem um contexto em comum, eles focalizam aspectos diferentes: o *frame* *Receive_visitor_scenario* (Cenário_da_recepção_do_visitante) reflete a perspectiva do anfitrião na recepção do visitante, e o *Visit_host* (Visitar_anfitrião) mostra, a partir do visitante, como se deu o contato com o anfitrião; ambos são perspectivas do *frame* *Visitor_and_host* (Visitante_e_anfitrião). Outra relação é Herança, na Figura 38, a seta vermelha indica que *Visiting_scenario_arrival*

(Cenário_da_visita_chegada) herda do *frame* *Arriving* (Chegar). Os EFs *Theme* (Tema) e *Goal* (Destino) do *frame* *Arriving* se tornarão os EFs *Agent* (Agente) e *Goal* (Destino) em *Visiting_scenario_arrival* (Cenário_da_visita_chegada), um é mais específico – no sentido de que um Agente pode ser pensado, nesse contexto, como alguém que se move, um Tema, por vontade própria –, e outro, igual, conforme o comportamento previsto pela FrameNet.

3.2.2 Cenário_do_Turismo

A atividade turística acontece em todas as partes do globo. A diversidade de culturas favorece a variedade de opções turísticas. Entretanto, ainda que se encontrem diferentes culturas e diferentes perfis de turista, o evento turístico é estruturado da mesma forma. Qualquer povo ao assumir o papel de turista precisa compreender o funcionamento das regras desse evento e ter ciência do papel a cumprir e, ainda que seja pela primeira vez, é capaz de fazer projeções acerca do que se deve esperar e fazer em cada etapa.

O estudo realizado para estruturar esse *frame*³¹ mostrou a necessidade de distingui-lo daqueles que tratavam do domínio da visita, já que o viés específico que o caracterizava se confrontava com o caráter amplo assumido pela visita e pela viagem, confirmando a necessidade de descrevê-lo em um *frame* próprio, ainda que estivesse evidenciada a existência de alguma relação entre esses domínios, até porque lexemas como *visitar* eram encontrados em ambas as situações.

- (42) [O ex-presidente _{AGENTE}] VISITOU ^{ALVO} [a mãe _{ENTIDADE}] [uma única vez _{FREQUÊNCIA}], [antes que ela deixasse o hospital Pró-cardíaco _{TEMPO}], [no Rio de Janeiro _{LOCAL}]. (Cetenfolha)
- (43) [Cerca de 3,5 milhões de pessoas _{AGENTE}] VISITARAM ^{ALVO} [ontem _{TEMPO}] [os 38 cemitérios de São Paulo _{ENTIDADE}]. O movimento, segundo o Serviço Funerário do Município, foi pelo menos duas vezes maior do que o do Dia de Finados do ano passado. (Cetenfolha)

³¹ Todos os *frames* modelados para o Cenário_do_turismo estão disponíveis em anexo a este trabalho. Além das análises ilustradas para esta seção, as demais podem ser visualizadas na guia Dados do site da FrameNet Brasil: <<http://www.framenetbr.ufff.br>>. Os padrões de valência estão em constante atualização uma vez que a anotação de sentenças nos *frames* que compõem o Cenário_do_turismo continua sendo feita diariamente.

Os exemplos (42) e (43),³² com a Unidade Lexical *visitar*, introduzem elementos que sugerem um tipo de visita que não se adéqua ao esperado para atividades turísticas. Naturalmente, visitas a hospitais e cemitérios não são os espaços ideais para o descanso ou o desfrute de estar diante de algo ou alguém. Barreto (1995), especialista dessa área, pondera a existência de diferentes tipos de viagens e afirma que não é possível tratar *viagem* e *turismo* como sinônimos.

Viagem não é a mesma coisa que turismo. O turismo inclui a viagem como uma parte, havendo muitas viagens que não são de turismo. Por exemplo, viagens de negócio, viagens de estudo, viagens para visitar parentes em condições especiais, como doença ou morte, podem ser, mais que um prazer, compromissos sociais (BARRETO, 1995, p. 13).

Ele ressalta que, ainda que compartilhem serviços em comum, as pessoas podem viajar por motivos alheios ao turismo. Todavia, isso não exclui o fato de se inserirem nesse *frame*, pois há vários casos de pessoas que acumulam atividades: ora se dedicam ao trabalho ou estudos, ora aproveitam para conhecer locais e atrações. Uma prova disso, como Barreto (1995) explicita, são congressos e outros eventos que inserem propostas turísticas na programação.

Diante de tal característica, a primeira tarefa foi especificar quais relações o *frame* *Cenário_do_turismo* mantinha com *Viajar* e *Cenário_da_visita* na tentativa de contemplar a relação com outros domínios. Enquanto o *Cenário_da_visita* inclui todos os tipos de visita, no domínio turístico, ela tem fim de entretenimento, o que evidencia que a visita turística equivale a um tipo específico de visita. Desse modo, seguindo a literatura, o *Cenário_do_turismo* é herdeiro do *Cenário_da_visita*, sendo que os Elementos de *Frame* do mais amplo são mantidos, de alguma forma, no *frame* mais específico. No *Cenário_da_visita*, há o EF Agente, correspondendo a quem pratica o ato de visitar, e o EF Entidade, lugar ou alguém que recebe a visita. Já o *Cenário_do_turismo* exhibe esses EFs como Turista e Atração, mais específicos que aqueles, por isso, diz-se que o *Cenário_do_turismo* é *frame* filho do *Cenário_da_visita*, que, por sua vez, é o mãe.

Quanto a *Viajar*, a relação não se dá da mesma forma. Se, por um lado, o *Cenário_do_turismo* não pode ser concebido como um tipo de viagem, por outro,

³² Essas sentenças foram extraídas do *corpus* Cetenfolha / Cetempúblico disponibilizado no Sketch Engine.

para compreender os *frames* estruturados para o turismo, tem-se o *frame* Viajar como *background*, indicando que os *frames* do Turismo são compreendidos quando se tem em mente o evento de viagem.³³ Nesse caso, a relação estabelecida é de Uso. A seta vermelha, que liga o Cenário_da_visita ao Cenário_do_turismo, Figura 39, especifica a relação Herança entre esses *frames*, enquanto a verde marca a de Uso com Viajar.

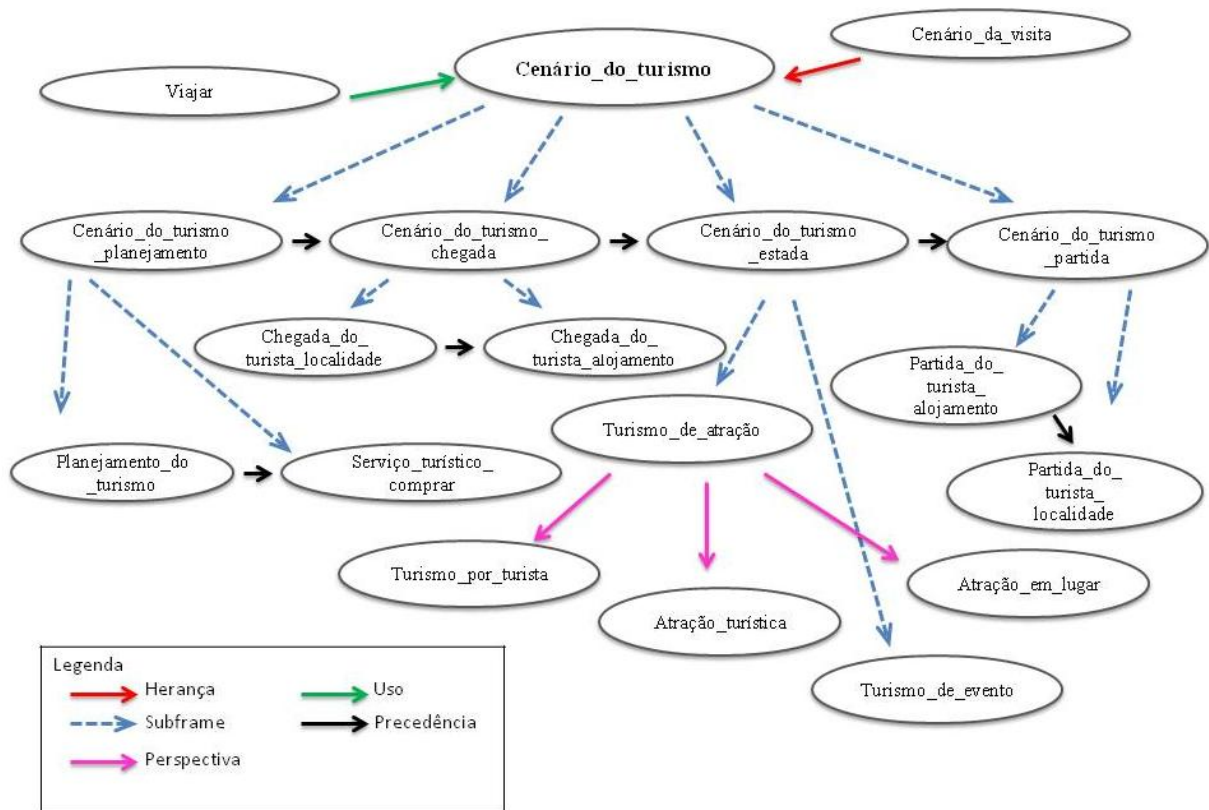


Figura 39: Cenário_do_turismo

A Unidade Lexical *turismo* é uma UL que representa o Cenário_do_turismo. Como os *frames* que dividem esse cenário descrevem eventos particulares, traçando perspectivas específicas, o Cenário_do_turismo é um *frame* não perspectivado. A seta azul indica a relação Subframe para as etapas

³³ Embora exista turismo no próprio local de moradia, o que significa que um habitante pode ser turista na sua própria cidade, sem necessariamente ter de viajar para visitar atrações turísticas, reconhece-se que a ideia da viagem esteja arraigada ao próprio conceito de turismo. Desse modo, a opção, neste trabalho, foi levar em conta o conceito prototípico de turismo legitimado pela sociedade, ainda que a literatura específica da área possa, corretamente, avaliar que isso seja mais um dos mitos consolidados sobre o Turismo. Caso se optasse por considerar essa avaliação, o Cenário_do_turismo usaria o *frame* Movimento_próprio (*Self_motion*), já que, na verdade, é necessário levar em conta o conceito de deslocamento para se conceber o de turismo.

desse cenário. Isso mostra que a atividade turística pode ser dividida em etapas, a partir de uma ordem temporal.

Primeiro, ocorre o *frame* `Cenário_do_turismo_planejamento`, dividido em dois subframes: `Planejamento_do_turismo` e `Serviço_turístico_comprar`. Estes *frames* descrevem a etapa anterior à chegada do turista ao destino. Por exemplo, o *frame* `Planejamento_do_turismo` insere questões como pesquisas para conhecer o local de destino, que indiquem a necessidade de vacinação, previsão do tempo, questões de documentação etc. O `Serviço_turístico_comprar`, por sua vez, detalha coisas como as compras que o turista realiza antes de chegar, como pacotes de turismo, incluindo deslocamento e passeios, por exemplo.

Após o *frame* do planejamento, há o `Cenário_do_turismo_chegada`, *frame* que trata da chegada do turista ao destino almejado. Este se divide em dois subframes, `Chegada_do_turista_localidade` e `Chegada_do_turista_alojamento`.

Segue-se à chegada do turista o `Cenário_do_turismo_estada`, *frame* que trata de acontecimentos e experiências durante o período da estada. Para este trabalho, foram apontados os subframes `Turismo_de_atração` e `Turismo_de_evento`. No `Turismo_de_atração`, três perspectivas são descritas nos *frames* `Turismo_por_turista`, `Atração_turística` e, por último, `Atração_em_lugar`. Já o `Turismo_de_evento` vai abordar os eventos que proporcionam atividades turísticas, como a Copa do Mundo. Tanto este *frame* quanto o `Cenário_do_turismo_planejamento` estão sendo desenvolvidos por outros mestrandos do projeto.

Por último, após a estada, tem-se o *frame* `Cenário_do_turismo_partida`, que descreve em dois subframes a `Partida_do_turista_alojamento` e a `Partida_do_turista_localidade`.

Embora não se possa dizer que sejam autônomos, os subframes que dividem o cenário turístico são eventos particulares, que descrevem situações específicas e complexas. Um turista pode não ter realizado um planejamento ideal da viagem com a efetuação de reservas etc, mas foi preciso, por exemplo, que comprasse passagens ou que realizasse a revisão do automóvel, eventos que ocorreram, necessariamente, antes da chegada ao destino. Do mesmo modo, não é possível

conceber o retorno do turista para o ponto inicial do deslocamento sem que, primeiramente, ele tenha participado do *frame* que trata da estada.

A função da seção seguinte é aprofundar os *frames* que modelam a chegada, a estada e a partida do turista do local turístico.

3.3 Frames da Experiência Turística

Como Leiper (1990 apud COPPER ET AL., 2002, p.38) avalia, o turista é o ator de todo esse sistema, por isso suas experiências delineiam a estrutura do `Cenário_do_turismo`.

3.3.1 Frames do `Cenário_do_Turismo_Chegada`

O *frame* `Cenário_do_turismo_chegada` é mais genérico do que os dois subframes que a ele se relacionam, no sentido de que aborda o evento da chegada do turista independentemente do fato de este ter chegado à localidade ou ao alojamento, sendo ambos os tipos de lugar agrupados no EF Destino. Isso faz dele um *frame* não perspectivado.

- (42) Para CHEGAR [à cidade DESTINO] [de ônibus MEIO_DE_TRANSPORTE], é preciso pegar as linhas que saem de Curitiba, Balneário Camboriú. [INI Turista] (Vai_Brasil)
- (43) As pessoas se falam, se cumprimentam, falam da vida, querem ouvir da sua. CHEGUEI [à Pousada Neusa Barbosa DESTINO] [tarde TEMPO]. Depois da longa viagem de BH até Tiradentes, que levou cerca de 4 horas, tive reunião de trabalho e cheguei para me instalar por volta da meia noite absolutamente cansada e faminta. [IND Turista] (Comi_perninha_de_cachorro)

As sentenças em (42) e (43) exemplificam esse *frame* com a Unidade Lexical *chegar*. A pesquisa por essa UL nos *corpora* mostrou que ela não diferencia, em valências sintáticas distintas, a chegada a uma cidade ou país da chegada a um alojamento, o que levou a postulá-la como evocadora do *frame* mais genérico. Vale ainda destacar a existência da relação *coreset* entre os EFs Turista e Meio_de_transporte, pois há casos cujo EF Meio_de_transporte é perfilado.

Os dois subframes desse cenário são a `Chegada_do_turista_localidade` e a `Chegada_do_turista_alojamento`.

No primeiro, as Unidades Lexicais *desembarcar* e *desembarque* são exemplos de ULs que evocam esse *frame*.

- (44) Fim das 6h pelo Delta, lindo pôr do sol antes de DESEMBARCAMOS [em Parnaíba DESTINO] [IND Turista] (Destino_de_viagem)
- (45) A pesquisa, que registrou [376 mil QUANTIDADE] DESEMBARQUES [em Búzios DESTINO] durante a temporada, mostra também que a maioria dos entrevistados(...). [INI Turista] (Brazil_tour)

Em (44), o EF Turista é uma Instanciação Nula Definida, o Destino, outro EF nuclear, é instanciado pelo sintagma *em Parnaíba*. Na sentença (45), a UL nominal *desembarque* tem o EF *Turista* como uma Instanciação Nula Indefinida, o Destino é representado pelo sintagma *em Búzios* e *376 mil* especifica a quantidade de desembarques realizados, EF periférico Quantidade.

Acerca do segundo subframe, o que trata da chegada ao alojamento, a sentença (46) o representa com a UL *check in*.

- (46) (...) seguir até a próxima cidade Campo Mourão ou seguir até Foz do Iguaçu onde teríamos que batalhar por vagas, pois [nosso TURISTA] CHECK IN estava programado só para o dia seguinte. A solução foi Campo Mourão. Chegamos por volta de 1 da manhã e o Paraná Palace Hotel foi a melhor possível, não que qualquer sofá não resolvesse, no entanto ficamos muito bem, ótimo café da manhã. [IND Alojamento] (Comi_perninha_de_cachorro)

O Quadro 4 resume os três *frames* criados para tratar da chegada do turista, apresentando as ULs evocadoras de cada um deles.

Frames da Experiência Turística	Tipo Semântico	Unidades Lexicais
Cenário_do_Turismo_Chegada	Não perspectivado	<i>chegar, chegada</i>
Chegada_do_Turista_Localidade	Perspectivado	<i>aportar, desembarcar, desembarque</i>
Chegada_do_Turista_Alojamento	Perspectivado	<i>check in, entrada, fazer check in</i>

Quadro 4: *Frames* da Experiência Turística – Chegada

3.3.2 *Frames* do Cenário_do_Turismo_Estada

O *frame* *Cenário_do_turismo_estada* trata do período em que o turista fica localizado no destino turístico, este *frame* de tipo semântico não perspectivado evoca Unidades Lexicais do tipo *estada*, *estadia* e *estar*.

Os dois subframes apontados neste trabalho foram o *Turismo_de_evento*, que especifica os tipos de evento que proporcionam atividades turísticas e o *Turismo_de_atração*. Sobre o *Turismo_de_atração*, decerto, é o *frame* que

assume papel proeminente no *Cenário_do_turismo_estada*. Ainda que haja diferentes tipos de turismo, como aqueles proporcionados por eventos, negócios e tantos outros, o *frame* *Turismo_de_atração* reúne todas as experiências proporcionadas pelas atividades turísticas. Assim, ainda que os torcedores que assistirão aos jogos da Copa do Mundo de 2014 estejam nas cidades brasileiras por conta de um evento específico, muitos deles também assumirão papel de turistas, uma vez irão participar das atividades englobadas no *Turismo_de_atração*.

Esse *frame* considerado de tipo semântico não perspectivado tem *atração* como Unidade Lexical. A partir de estudo de *corpora*, percebeu-se a existência de três perspectivas para tratar do *frame* *Turismo_de_atração*: a do turista, a da atração turística e a do lugar.

Layer	O	t	u	r	i	s	t	a	p	o	d	e	c	o	n	h	e	c	e	r	a	s	b	e	l	e	z	a	s	l	o	c	a	i	s	t	a	n	t	o	d	u	r	a	n	t	e				
FE	T	u	r	i	s	t	a															A	t	r	a	ç	ã	o															T	e	m	p	o				
GF_br	E	x	t	r	a																	O	b	j	e	c	t	o																		D	e	p			
PT_br	N	P																					N	P																											
Outros...																																																			

Figura 40: Sentença com UL *conhecer*

Layer	O	J	a	l	a	p	ã	o	a	p	r	e	s	e	n	t	a	u	m	p	a	n	o	r	a	m	a	d	e	p	e	r	d	e	r	o	f	ô	l	e	g	o								
FE	L	u	g	a	r																	A	t	r	a	ç	ã	o																						
GF_br	E	x	t	r	a																		O	b	j	e	c	t	o																					
PT_br	N	P																					N	P																										
Outros...																																																		

Figura 41: Sentença com UL *apresentar*

Layer	A	r	e	g	i	ã	o	a	i	n	d	a	a	b	r	i	g	a	g	r	u	t	a	s	d	e	g	r	a	n	d	e	b	e	l	e	z	a												
FE	L	u	g	a	r																	A	t	r	a	ç	ã	o																						
GF_br	E	x	t	r	a																		O	b	j	e	c	t	o																					
PT_br	N	P																					N	P																										
Outros...																																																		

Figura 42: Sentença com UL *abrigar*

Veja que as sentenças ilustradas pelas Figuras 40, 41 e 42 perfilam elementos diferentes. Na Figura 40, a sentença “o turista pode conhecer as belezas locais tanto durante o dia quanto durante a noite” assume a perspectiva do turista para apresentar essa experiência. Já, em 41, “o Jalapão apresenta panorama de perder o fôlego”, o foco é a atração turística, e, na sentença da Figura 42, “a região ainda abriga grutas de grande beleza”, a perspectiva assumida é a do lugar que possui ou hospeda alguma atração. Convencionou-se caracterizar cada uma dessas perspectivas, respectivamente, como *Turismo_por_turista*, *Atração_turística* e *Atração_em_lugar*.

O estudo desses *frames*, a partir de seus elementos nucleares, possibilita algumas generalizações. Em *Turismo_por_turista* e *Atração_turística*, os EFs nucleares são Turista, Atração e Lugar, sendo que os dois últimos atuam em relação de *coreset*, o que significa que ambos são nucleares, mas a presença de um satisfaz a valência semântica e sintática do predicador. O comportamento desses EFs nos *corpora* mostrou que a posição de cada um em relação ao predicador – no caso, os verbais – implicava enfatizar papéis semânticos específicos. Desse modo, na perspectiva trazida pelo *frame* *Turismo_por_turista*, a valência das Unidades Lexicais verbais era preenchida por dois argumentos, o que assumia função gramatical Externo era o EF Turista, e aquele de função gramatical Objeto Direto era Atração ou Lugar. O *frame* *Atração_turística*, por sua vez, perfilava o EF Atração ou Lugar como Externo, e o EF Turista era instanciado com função de Objeto Direto ou Indireto.

- (47) Para que [os turistas Turista] CONHEÇAM [as belezas naturais ATRAÇÃO], são organizados passeios arqueológicos na mata, nos rios e nas praias. (Vai_Brasil)
- (48) [A Praia da Gamboa ATRAÇÃO] OFERECE [belezas naturais Descrição] e atrai surfistas de diversas partes do Brasil. [IND Turista](Vai_Brasil)

As sentenças (47) e (48) exemplificam essa observação. A UL *conhecer*, em (47), representa um esquema com dois argumentos, um ocupado pelo EF Turista na posição de Externo, e o outro, pelo EF Atração, na posição de Objeto Direto. Em (48), a Unidade Lexical *oferecer* apresenta o EF Atração, a *Praia de Gamboa*, com função gramatical Externo, o EF Descrição, *belezas naturais*, atua com FG Objeto Direto, e o EF Turista é uma Instanciação Nula Definida.

A terceira perspectiva encontrada para o *frame* *Turismo_de_atração* foi *Atração_em_lugar*. Nesse *frame*, o EF Turista não assume papel nuclear, em muitos casos, sequer é mencionado. A Unidade Lexical *possuir*, na sentença (49), requer dois argumentos, um preenchido pelo EF Lugar, e o outro, pelo EF Atração. O foco é especificar o potencial de um lugar em possuir ou hospedar recursos naturais ou artificiais reconhecidos pela humanidade, e não necessariamente no turista.

- (49) [A cidade LUGAR] POSSUI [uma grande variedade de belezas naturais, formando um ecossistema único no Estado ATRAÇÃO].

O Quadro 5 resume os *frames* criados para tratar da estada do turista, apresentando as ULs que evocam cada um deles.

<i>Frames da Experiência Turística</i>	<i>Tipo Semântico</i>	<i>Unidades Lexicais</i>
<code>Cenário_do_Turismo_Estada</code>	Não perspectivado	<i>estada, estadia, estar</i>
<code>Turismo_de_Atração</code>	Não perspectivado	<i>atração</i>
<code>Turismo_por_Turista</code>	Perspectivado	<i>visitar, conhecer, turista, visitante, desfrutar, apreciar, fazer tour</i>
<code>Atração_Turística</code>	Perspectivado	<i>oferecer, atrair, surpreender, apresentar, destacar-se</i>
<code>Atração_em_Lugar</code>	Perspectivado	<i>abrigar, possuir, ter, haver</i>

Quadro 5: *Frames da Experiência Turística – Estada*

3.3.3 *Frames do Cenário_do_Turismo_Partida*

Após as atividades turísticas terem sido realizadas, o turista finaliza a estada e parte para o destino, que, geralmente, é o ponto inicial do deslocamento. O *frame* `Cenário_do_turismo_partida` de tipo semântico não perspectivado é representado por ULs como *deixar* e *partir*.

- (50) Se quiser DEIXAR [a Baía de todos os Santos ORIGEM] [pela barra falsa TRAJETÓRIA], lembre-se: seu barco precisa ter menos de 19m de altura, da linha d' água ao topo do mastro, por causa da ponte que existe ali. [CNI Turista] (Brazil_tour)

O exemplo (50) ilustra esse cenário com a Unidade Lexical *deixar*. Note que o EF Turista é uma Instanciação Nula Construcional, uma vez que é um discurso instrucional. Os dois subframes destacados para representar esse cenário foram `Partida_do_turista_alojamento` e `Partida_do_turista_lugar`. Eles se ligam pela relação de precedência, uma vez que, primeiramente, o turista finaliza o contrato ou acordo estabelecido com o alojamento para depois partir da localidade. As sentenças em (51) e (52) ilustram esses frames.

- (51) Acesso *free* ao Business Center, Room Service 24 horas, CHECK OUT inteligente. Considerado o melhor hotel da cidade [INI Turista] [DNI Alojamento] (Fifa)
- (52) Para isso, é preciso reunir e apresentar na hora do embarque ou do CKECK IN a documentação e autorizações necessárias. [INI Turista] [INI Destino] (Brasil_Gov)

Em (51), a Unidade Lexical *check out* destaca a partida do hóspede do local de hospedagem. O EF Turista é uma Instanciação Nula Indefinida, pois não se tem

pelo contexto a especificação de que se trata de um turista, já que a sentença advém do *corpus* FIFA. O EF Alojamento aparece como Nula Definida, uma vez que a informação de que se trata de um hotel é especificada no contexto. Já a sentença em (52), especifica com a UL *check in* a entrada do Turista no local onde fará o embarque e, conseqüentemente, a saída da localidade.

Assim como nos *frames* que tratam da chegada, os da partida também podem perfilar o EF Meio_de_transporte. A sentença (53) mostra essa ocorrência. O alvo *sair* coloca em foco o Meio_de_transporte *ônibus*.

(53) [De Petrolina e de Teresina _{LUGAR}], SAEM^{ALVO} [ônibus _{MEIO_DE_TRANSPORTE}] [para São Raimundo Nonato _{DESTINO}]. (Vai_Brasil)

O Quadro 6 resume os *frames* criados para tratar da partida do turista e apresenta as ULs que evocam cada um deles.

Frames da Experiência Turística	Tipo Semântico	Unidades Lexicais
Cenário_do_Turismo_Partida	Não perspectivado	partir, deixar, saída
Partida_do_Turista_Alojamento	Perspectivado	<i>check out</i> ,
Partida_do_Turista_Localidade	Perspectivado	<i>check in</i> , embarque, embarcar

Quadro 6: *Frames* da Experiência Turística – Partida

4 CONTRIBUIÇÕES DA MODELAGEM DO CENÁRIO DO TURISMO PARA O DESENVOLVIMENTO DO COPA 2014

Estruturar o *Cenário_do_turismo* permitiu fazer algumas contribuições teóricas sobre o desenvolvimento de dicionários eletrônicos multilíngues. Dentre elas estão a utilização de frames como interlíngua, a reflexão sobre como o Kicktionary une os conceitos de *frame* e *synset* na composição da sua proposta de dicionário e a decorrente proposta de revisão do Kicktionary e do papel da WordNet no Copa 2014. A ideia é poder utilizar as duas teorias linguísticas explorando aquilo que cada uma delas pode oferecer de melhor.

4.1 Frames como Interlíngua

Ainda que as máquinas estejam alcançando níveis consideráveis de inteligência, é dado que, quanto maior o nível de abstração do *input* solicitado, maior é a limitação que a máquina enfrentará no processamento de seus dados. Isso é corroborado em Di Felippo e Dias da Silva (2009), que assinalam que, se colocado em uma escala de abstração e complexidade, o conhecimento linguístico pode ser hierarquizado da seguinte forma: pragmática > semântica > sintaxe > morfologia. A partir dessa escala, compreendem-se os motivos que levam as máquinas a responderem, razoavelmente bem, àquelas informações que podem ser depreendidas por conhecimento morfológico e sintático. Entretanto, quando é necessário tratar informações que exijam conhecimento semântico e/ou pragmático, as pessoas se deparam com tradutores automáticos e sistemas de busca “pouco inteligentes”.

Para os dicionários multilíngues, esse desafio é reiterado. Um propósito básico de um recurso lexical desse tipo é ser capaz de facilitar a compreensão de sistemas linguísticos diferentes. Devido a esse intuito, modelar um dicionário multilíngue em meio eletrônico demanda especificidades que não são requeridas pelos materiais impressos de mesma finalidade, pois é importante que o sistema computacional onde esses materiais estejam armazenados possa se comunicar na busca por traduções.

Dias da Silva et al. (2007) mostram, motivados pelos estudos de Dorr et al. (2000 apud DIAS-DA-SILVA ET AL., 2007), que os procedimentos para a realização da tradução automática podem ser classificados de acordo com o paradigma adotado e o método escolhido. O paradigma diz respeito aos componentes de representação do conhecimento, e os métodos podem ser tradução direta, transferência ou interlíngua, Figura 43.

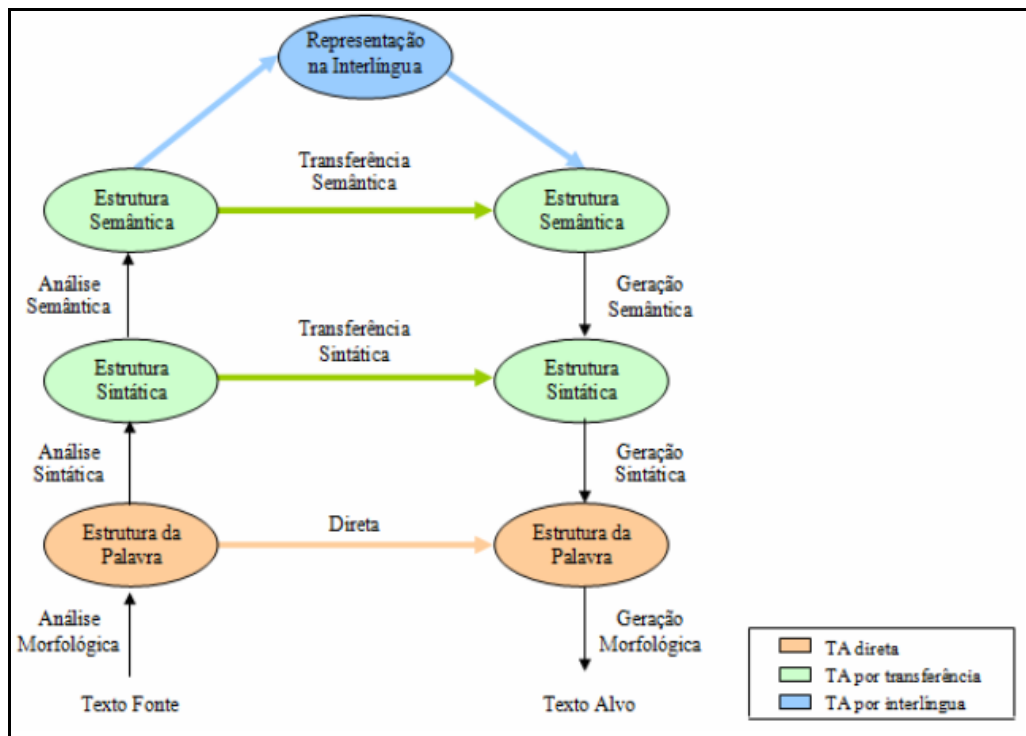


Figura 43: Níveis de profundidade do conhecimento nos sistemas de tradução automática
Fonte: Dorr et al. (2000 apud DIAS-DA-SILVA ET AL., 2007, p. 62)

A tradução direta é o método mais simples, pois utiliza o mínimo possível de processamento linguístico. O mais comum nesse método, como os autores demonstram, é o contraste de dois sistemas linguísticos, pois possibilita o uso de um dicionário bilíngue para viabilizar a tradução automática, que, geralmente, acontece via palavra. Dias da Silva et al. (2007) elencam várias dificuldades proporcionadas por esse tipo de tradução: o banco de dados pode não ter o equivalente buscado, pode não reconhecer as construções gramaticais utilizadas ou mesmo a estrutura dos constituintes, problemas que influenciarão na qualidade do produto final.

Em contraste com o método de tradução direta, os métodos indiretos proporcionam análises mais refinadas, pois, conforme os autores evidenciam, reconhece-se a necessidade de entender o sistema linguístico e também

extralinguístico das línguas envolvidas. No método indireto de transferência, Dias da Silva et al. (2007) destacam as etapas que comumente são seguidas, como a necessidade de criar uma representação intermediária na língua fonte para se comunicar com a representação intermediária da língua alvo e, assim, gerar a sentença nessa língua. Na Figura 36, são ilustrados esses processos que marcam a fase de análise da língua fonte, a fase de transferência de dados de uma língua para outra e a fase de geração de dados da língua alvo. Por esse método, é possível, conforme explicitado pela Figura 36, transferir tanto informações sintáticas quanto semânticas.

Todavia, com a dificuldade de conseguir criar as regras necessárias para transferir dados linguísticos, surge o método indireto de interlíngua, que, na avaliação dos autores, tem como objetivo capturar o significado a ser transmitido, independentemente da língua a ser traduzida. A interlíngua, nesse sentido, é um sistema linguístico de representação, responsável por realizar o elo de comunicação entre as línguas. De acordo com Dias da Silva et al. (2007), o método de interlíngua consiste na análise completa do texto na língua fonte para que o significado seja extraído e representado na interlíngua. Após a realização desse processo, ocorre a geração do texto na língua desejada. Não há aqui o objetivo de encontrar equivalências estritas na língua alvo, o que, na avaliação dos autores, torna a interlíngua mais próxima de uma paráfrase do que de uma tradução, propriamente dita, pois não se tem garantia, nem pretensão, de que o texto fonte seja transposto no texto alvo.

Esses métodos podem ser estruturados de modo unidirecional ou bidirecional. Se tiver apenas uma direção, “uma língua é fonte ou alvo”, mas não as duas coisas, se for bidirecional, isso significa que “pode ocorrer tanto de L1 para L2, quanto de L2 para L1” (cf. DIAS-DA-SILVA ET AL., 2007, p.68). Juntamente com os métodos de tradução automática, diferentes paradigmas podem ser seguidos. Arnold et al. (1993 apud DIAS-DA-SILVA ET AL., 2007, p. 68-74) discutem dois tipos de paradigmas, o fundamental e o empírico. Os modelos de paradigma fundamental conduzem a tradução automática com teorias linguísticas bem definidas. Há diversas possibilidades, como o modelo estruturado por léxico e por conhecimento, por exemplo. No modelo de tradução automática baseada em léxico, existem regras para efetuar as ligações entre os itens lexicais de cada língua. Já o estruturado por conhecimento tem como objetivo fornecer, além de dados linguísticos, informações

extralinguísticas para que o sistema seja capaz de efetuar inferências. Em contraste, os modelos empíricos abrem mão de teorias linguísticas para empregar técnicas, como as baseadas em estatísticas. Há de se ressaltar, conforme Dias da Silva et al. (2007) assinalam, a existência de sistemas que optam por mesclar diferentes tratamentos, os chamados paradigmas híbridos.

Pela visualização desse panorama teórico, pretende-se, com o dicionário Copa 2014, explorar o potencial da Semântica de *Frames* e analisar em que medida esse paradigma linguístico, que seria classificado, nos estudos de Arnold et al. (1993 apud DIAS-DA-SILVA ET AL.. 2007, p. 68-74), como paradigma fundamental baseado em conhecimento, pode ser útil no processo de tradução automática. Segundo a análise de Fillmore, as relações como sinonímia, antonímia e demais relações lexicais não são comparáveis àquilo que as propriedades semânticas podem oferecer se vinculadas a *frames* (FILLMORE & CALLEJAS, 2003, p.45). Isso porque o trabalho com *frames* possibilita recuperar noções semânticas importantes, que não são possíveis através de relações lexicais, como a relação da palavra *turista* com *check-in*, por exemplo. Dessa maneira, assumindo que é imperativa a necessidade de bancos de dados lexicais conseguirem transmitir informações semânticas às máquinas, a proposta é investigar como os *frames* atuariam se utilizados como interlíngua.

Em Lönneker-Rodman (2007), é apresentado como o estágio atual da base de dados da FrameNet lida com a multilingualidade. O autor insere discussões teóricas que especificam onde, atualmente, essa rede semântica está localizada quando a temática é o desenvolvimento de recursos multilíngues. Para tal, o ponto de referência utilizado é a EuroWordNet (Vossen, 2002), iniciativa multilíngue ilustrada na seção 2.3 deste trabalho, que emprega o banco de dados da WordNet de Princeton como índice interlingual no alinhamento dos *synsets* das diferentes wordnets. Em linhas gerais, o autor mostra que as bases de dados estruturadas em torno da FrameNet apresentam o mesmo princípio organizacional, já que a metodologia seguida é a mesma. Porém, se cada língua for representada de acordo com suas especificidades, a inter-relação dos dados não apresentará equivalência, indicando que as informações linguísticas não irão se alinhar perfeitamente. Isso acontece porque os *frames* em línguas diferentes não necessariamente são representados da mesma forma. Além disso, a polissemia não é um fenômeno uniforme, ou seja, uma palavra pode ser polissêmica em uma língua, mas seu

equivalente na outra pode não ter o mesmo status de polissemia³⁴, o que será refletido pelas Unidades Lexicais, que podem ainda não encontrar um correspondente na outra língua.

Tendo em vista essa realidade, Lönneker-Rodman (2007) expõe duas metodologias discutidas pela EuroWordNet para relacionar bancos de dados multilíngues: o método de fusão e o de expansão. O método de fusão é o que acontece quando os bancos de dados são construídos de forma independente. Já o método de expansão usa o banco de dados de uma língua como modelo para incluir os dados de outra língua. O autor bem assinala que os dois métodos apresentam problemas: o de fusão acarreta dificuldades no processo de alinhamento de dados, já que os sistemas linguísticos marcam especificidades no processo de lexicalização dos conceitos, e o método de expansão negligencia essas especificidades quando analisa seus dados se limitando à sistematização dos dados da língua modelo.

Para a construção do dicionário Copa 2014, a ideia inicial era realizar uma parceria entre a FrameNet de Berkeley e a Spanish FrameNet para que os dados do domínio turístico fossem analisados através de suas próprias plataformas. Entretanto, como o *Cenário_do_turismo* não havia, até então, sido desenvolvido por esses grupos, o prazo estabelecido para a estruturação do dicionário impediu que se unissem a esse empreendimento. Dessa forma, a alternativa foi estruturar o *Cenário_do_turismo* também para as línguas inglesa e espanhola.

Pelas considerações de Lönneker-Rodman (2007), pode-se dizer que o Copa 2014, em si, constitui-se como um híbrido dos dois métodos: fusão e expansão. Esse tratamento acontece, porque, como demonstrado a seguir, os *frames* dos domínios desenvolvidos pelo dicionário não apresentam divergências socioculturais. Como o *Cenário_do_turismo* e os *frames* do domínio da Copa do Mundo não foram encontrados na base de dados da FrameNet mãe, a alternativa foi estruturá-los, o que aconteceu primeiramente para a língua portuguesa³⁵. Num segundo momento, foi possível partir desses *frames* para a criação dos demais – método de expansão. Porém, isso não impedia que os *frames* fossem estruturados de modo independente, pois não foi encontrada divergência quanto à estruturação desses eventos – método de fusão.

³⁴ Ver exemplo da Unidade Lexical *marcar* na seção 2.1 deste trabalho.

³⁵ Essa realidade não pode ser estendida para os *frames* do domínio do futebol. A equipe liderada pela professora Rove Chishman, responsável por este domínio, pôde partir dos frames do inglês, previamente construídos no Kicktionary.

Para realizar esse alinhamento, a opção foi criar uma relação de correspondência entre *frames*. Isso significa que os *frames* do português estarão ligados aos das demais línguas por tal relação. Em linhas gerais, essa relação se estrutura de modo familiar à relação de Herança da FrameNet. Como não houve divergência, a estrutura dos *frames*, a partir de seus elementos, será alinhada entre os bancos de dados.

Isso foi possível porque o estudo realizado aponta que o evento turístico é perspectivado da mesma forma pelas diferentes culturas e lugares do mundo. Desse modo, propõe-se que a pesquisa com os dados do português do Brasil para a definição do cenário sirva também para representar o cenário turístico do inglês e do espanhol. Embora haja diferentes culturas e sistemas econômicos, o turismo é uma atividade global, que acontece por todas as partes do mundo, gerando experiências, comportamentos e atitudes que acontecem dentro de um mesmo evento padrão. Comportamento semelhante é percebido com o evento mundial da FIFA. A modalidade de futebol da Copa do Mundo padroniza regras universais para todos os países. Dessa forma, ainda que sejam vistas Unidades Lexicais específicas para lexicalizar determinado conceito tanto no domínio turístico quanto esportivo, a estrutura dos *frames* quanto à nuclearidade dos EFs e relações entre *frames*, por exemplo, não se distinguirá.

Todavia, por mais que este dicionário não tenha de lidar diretamente com a dificuldade de alinhamento de dados, essa discussão precisa ser realizada para propor a utilização dos *frames* como interlíngua, já que essa proposta vem gerando discussões em torno da FrameNet e do tratamento de recursos lexicais multilíngues de modo geral. O trabalho realizado por Bertoldi (2011) é uma demonstração dos vários desafios ao se lidar com *frames* para representar diferentes línguas. O autor analisa em que medida os *frames* do inglês podem ser utilizados para o português do Brasil no que diz respeito ao sistema jurídico. Em estudo contrastivo, a conclusão alcançada foi de que o melhor a se fazer é a estruturação de *frames* específicos para representar o sistema jurídico brasileiro, visando a um posterior alinhamento de dados. Como Bertoldi avalia, a jurisdição estadunidense se baseia na *common law* enquanto a brasileira é fundamentada na *civil law*, sistemas jurídicos herdados de culturas distintas, que, hoje, representam culturas também diferentes. Ainda que sejam encontrados aspectos considerados semelhantes nesses sistemas, as particularidades devem ser descritas em *frames* distintos.

Procurando uma classificação que atendesse a essas especificidades, o autor destaca que distinções como *frames* universais e *frames* dependentes de língua não seriam ideais e não teriam respaldo teórico (cf. BERTOLDI, 2011, p. 110-117). Como Fillmore (1982, p.112) já dissera, “as palavras representam categorizações de experiências, e cada uma dessas categorias se baseia em uma situação motivadora, que ocorre segundo um contexto de conhecimento e experiência”³⁶. Desse modo, como muito bem ressalta Bertoldi (2011), o *frame* é culturalmente orientado. Estendendo essa discussão, a distinção que Fillmore (1985) faz de *frames* inatos e *frames* aprendidos também não deveria ser considerada categoricamente, haja vista que tanto os *frames* inerentes ao desenvolvimento cognitivo (advindos da percepção dos sentidos, por exemplo) quanto aqueles estruturados pela relação do ser humano com o mundo (vide a atividade turística e o sistema jurídico) sofrem interferência sociocultural.

Ante a essa realidade, a proposta de Bertoldi é rever a aplicabilidade de *frames* como interlíngua, pois, por conterem aspectos culturais, em muitos casos, os *frames* não são equiparáveis, o que ficou comprovado pelos *frames* do domínio jurídico. Entretanto, é exatamente essa conclusão que necessita ser questionada. Os *frames* são a oportunidade de representar a modelagem dos eventos, situações e experiências vivenciadas pelas sociedades. Com essa teoria, pautada pela diferença, não se concebe a tentativa de apagar as divergências socioculturais, mas também não se exclui o fato de que um dado evento ou experiência seja descrito da mesma forma em línguas distintas.

A plataforma da FrameNet, de fato, não consegue responder completamente aos diversos desafios colocados para a constituição de recursos lexicais tanto monolíngues quanto multilíngues. Isso acontece por diferentes motivos, a saber:

- i. O banco de dados da FrameNet está em constante atualização. Desse modo, as análises não estão esgotadas. Como não há todos os *frames*, as relações entre *frames* são passíveis de questionamentos ou adaptações, é o que se vê hoje na relação *Uso* entre os *frames* *Visiting* e *Touring*. Na verdade, conforme demonstrado no capítulo 3, ao aprofundar-se o estudo sobre esse domínio, percebeu-se que se trata de uma relação de Herança. Porém, deve-se levar em consideração que

³⁶ “(...) words represent categorizations of experience, and each of these categories is underlain by motivating situation occurring against a background of knowledge and experience”

questões como essas são inerentes ao fato de ser um projeto em andamento.

- ii. O intuito de modelar a linguagem em uso é uma tarefa de muitos desafios, o que traz a necessidade de rever constantemente a teoria para sistematizar a prática, sem correr o risco de negligenciar a realidade.
- iii. Como Lönneker-Rodman (2007) assinalou nas conclusões de suas avaliações sobre a FrameNet como recurso multilíngue, hoje, as condições de equivalência entre línguas não estão formalmente definidas. O motivo disso é que a FrameNet foi estruturada de acordo com os dados do inglês. As definições dos *frames* bem como as realizações dos EFs e padrões sintáticos foram formalizadas para essa língua. Isso não impede que essa rede semântica se expanda para outras línguas e nem que seja utilizada como índice interlingual, mas aponta a necessidade de revisão analítica e teórica. Além disso, as relações entre *frames* foram estruturadas para retratar os *frames* internos a um sistema linguístico, e não para realizar o mapeamento entre línguas.

Atualmente, um recurso da FrameNet vem sendo explorado como iniciativa multilíngue – o Frame SQL. Esse recurso desenvolvido por Sato (2003, 2008) é um sistema de busca com várias funcionalidades para que o usuário possa pesquisar o banco de dados da FN com um nível maior de detalhamento. Além de ter a opção de busca por Unidades Lexicais, o usuário também especifica as propriedades sintáticas desejadas em cada Elemento de *Frame*. Além da base de dados do inglês, algumas *framenets*, como a Spanish FrameNet e a Japanese FrameNet, estão explorando esse recurso para contrastar seus bancos de dados.

A realização dos elementos de *frame* da Unidade Lexical *aproximarse* do banco de dados da Spanish FrameNet e *approach* do Inglês é contrastada por essa ferramenta, ilustrada pelo Quadro 7. Os números representam quantas sentenças foram encontradas com cada padrão descrito. Há padrões comuns às duas línguas e também há combinações específicas. A proximidade ou não entre os sistemas linguísticos é evidenciada nos padrões de realização. Entretanto, uma ressalva a se fazer é que, para encontrar as possíveis traduções, essa ferramenta utiliza um dicionário bilíngue *online*, no caso, espanhol-inglês, que retorna a busca com os sinônimos descritos nos verbetes do dicionário. A função do *frame* bem como de todo o estudo empírico realizado pela análise da FrameNet é colocada de lado, já

que a função é mostrar quais palavras listadas nos verbetes do dicionário utilizado aparecem na lista de Unidades Lexicais de cada língua. Dessa forma, o potencial do *frame* como interlíngua não é plenamente utilizado.

	aproximarse	approach
Goal + Manner + Theme		01
(Goal) + Manner + Theme	01	01
Depictive + Goal + Place + Theme	01	
Depictive + Goal + Manner + (Theme)	01	
(Goal) + Source + Theme		04
(Goal) + Manner + (Theme)		01
Degree + Goal + Theme	01	
Goal + (Theme)	02	
Goal + Theme	07	12
Goal + Theme + Time	01	
Goal + Means + Theme		01
(Goal) + Result + Theme	01	
Goal + Manner + Manner_of_motion + Theme	01	
Goal + Path + Theme		05
Depictive + (Goal) + Manner + Theme	01	
Cotheme + Goal + Theme	01	
Goal + Source + Theme	01	06
Circumstances + (Goal) + Manner + (Theme)	01	
(Goal) + Path + Theme		01
Goal + Speed + Theme + Time	01	
(Goal) + Theme		02

Quadro 7: Realização semântica de *aproximarse* e *approach*³⁷
 Fonte: SATO, 2008, p. 760

Para o âmbito lexicográfico e computacional, várias medidas podem ser tomadas para aperfeiçoar essa rede semântica:

- i. um tratamento mais detalhado do tipo semântico de cada Elemento de *Frame*, alternativa proposta por Bertoldi (2011) para realizar o contraste entre diferentes sistemas jurídicos;
- ii. cada definição de *frame* é resultado de um detalhado estudo de *corpora*,³⁸ que fornece evidências dos padrões não só semânticos como também sintáticos. Como esse levantamento indica regularidades na realização dos *frames*, a valência sintática, além da semântica, também pode ser empregada para a análise entre línguas;

³⁷ Os elementos de *frame* entre parênteses marcam uma instanciação nula.

³⁸ É preciso ponderar que, para o domínio do turismo no dicionário Copa 2014, optou-se por textos comparáveis, isto é, fontes que tratam de domínios textuais semelhantes, produzidos pelas próprias comunidades linguísticas. Nesse ponto, diverge-se da concepção vista em Boas (2009) e Bertoldi (2011), que avaliam que apenas os textos traduzidos ou paralelos fornecem evidências necessárias para contrastar línguas.

- iii. as relações entre *frames* e entre Elementos de *Frames* precisam ser adaptadas para a investigação multilíngue. Relações como a de herança podem considerar um tipo de herança especial, com o objetivo de mapear as semelhanças e particularidades de cada língua ao perspectivar um dado conceito;
- iv. um tratamento refinado das Unidades Lexicais tanto no sentido de relacioná-las por categorias de palavras lexicais, semelhante ao que é realizado hoje pela WordNet, quanto no sentido de organizar a polissemia aperfeiçoaria o banco de dados para a prática lexicográfica. A plataforma precisa, primeiramente, sistematizar os usos polissêmicos em banco de dados monolíngue para depois seguir um propósito multilíngue. Para isso, uma possibilidade é pesquisar, baseando-se em estudo de *corpora*, o comportamento de lexemas polissêmicos na tentativa de mapear a proximidade semântica entre esses itens lexicais e sistematizar a representatividade de cada uso.

Por mais que ainda não se tenham respostas para os desafios lexicográficos enfrentados pelo banco de dados da FrameNet, o caráter transcultural dos domínios abordados pelo dicionário permite que os *frames* sejam utilizados como interlíngua na estruturação do Copa 2014.

4.2 Proposta de Revisão da Estruturação do Kicktionary

A modelagem do cenário do Turismo para o dicionário Copa 2014 também permite fazer algumas reflexões teóricas acerca do método seguido por Schmidt (2006, 2007, 2008, 2009) no Kicktionary. Ainda que o autor pondere que a criação do dicionário se dirige à lexicografia computacional focada para usuários humanos, e não para o aprendizado de máquinas, é pertinente levantar algumas questões a respeito dos conceitos que guiam a elaboração desse recurso.

A organização do evento futebolístico baseada na Semântica de *Frames* permite relacionar e sistematizar a maneira como as informações linguísticas estão associadas ao conhecimento de mundo, e isso, como autor destaca, não é contemplado de modo eficaz pelos dicionários tradicionais. Contudo, para estruturar

essas questões na interface do dicionário, o autor retoma a distinção feita por Fillmore no texto de 1977,

eu quero dizer que as pessoas, na aprendizagem de uma língua, vêm a associar determinadas cenas com determinados *frames* linguísticos. Pretendo usar a palavra *cena* – uma palavra com a qual eu não estou completamente satisfeito – em um sentido geral ao máximo, para incluir não apenas cenas visuais, mas os tipos familiares de relações interpessoais, cenários padrão, layouts familiares, estruturas institucionais, experiências ativas, imagem corporal e, em geral, qualquer tipo de segmento coerente, grande ou pequeno, crenças humanas, ações, experiências, ou imaginação. Eu pretendo usar a palavra *frame* para me referir a qualquer sistema de escolhas linguísticas – os casos mais simples sendo coleções de palavras, mas incluindo também as escolhas de categorias gramaticais – que pode ser associado com instâncias prototípicas das cenas. (FILLMORE, 1977, p. 82)³⁹

Quando Fillmore fez essa distinção, o propósito era discernir estruturas de conhecimento mais amplas e abstratas daquelas calcadas nas escolhas linguísticas dos falantes para representar determinado conhecimento. Entretanto, ao fazer isso, Fillmore dá margem a uma dicotomia que já havia sido superada no meio acadêmico, que é a dissociação entre conhecimento linguístico e conhecimento enciclopédico. Com isso, no texto *Frame Semantics* de 1982, obra introdutória à Semântica de *Frames*, ele amplia o conceito de *frame* para “qualquer sistema de conceitos relacionados de tal forma que, para entender qualquer um deles, você tem de entender toda a estrutura na qual ele se encaixa”. Ao abranger esse conceito, o autor busca mostrar que todas as distinções outrora realizadas para especificar características peculiares observadas nos *frames* se atrelam a um único conceito.

Há um uso não linguístico comum desta palavra [*frame*], que não se refere ao que uma porção de linguagem evoca na mente de um intérprete, mas aos tipos de estruturas conceptuais que um intérprete invoca para construir o sentido de alguma experiência. Este uso não linguístico do conceito tem uma longa história (BARTLETT, 1932; PIAGET, 1971; MILL, 1846; cf. FILLMORE, 1985), mas floresceu nas ciências cognitivas e sociais na década de 1970. A ideia por trás deste trabalho é que as pessoas entendam o que estão observando ao recrutar memórias de experiências passadas, ou ao construir o que é observado como instâncias ou variações de estruturas de crença e experiência, que poderiam ser utilizadas para entender o que se tem ouvido ou observado. Estas estruturas da

³⁹ I want to say that people, in learning a language, come to associate certain scenes with certain linguistic frames. I intend to use the word scene – a word I am not completely happy with – in a maximally general sense, to include not only visual scenes but familiar kinds of interpersonal transactions, standard scenarios, familiar layouts, institutional structures, enactive experiences, body image; and, in general, any kind of coherent segment, large or small, of human beliefs, actions, experiences, or imaginings. I intend to use the word frame for referring to any system of linguistic choices – the easiest cases being collections of words, but also including choices of grammatical rules or grammatical categories – that can get associated with prototypical instances of scenes.

mente são chamadas *frames* e não estão, necessariamente, conectadas a formas específicas de linguagem. (FILLMORE, 2010, p. 614)⁴⁰

Para Schmidt, a noção de cenas e *frames* fornece um método sistemático para marcar a hierarquia dessas estruturas. Por ser de natureza independente da linguagem, as cenas atuam de forma supeordenada ao *frame* e assim poderiam agrupar *frames* em diferentes línguas, respeitando as especificidades de cada sistema linguístico ao lexicalizar um dado conceito, daí, a justificativa do autor ao assinalar a eficácia desses conceitos para recursos multilíngues. A cena vai, então, no Kicktionary, designar eventos prototípicos do futebol, que reunirão diversos *frames* relacionados.

Por mais que estruturar a partida de futebol imponha especificidades diante da estruturação da atividade turística, o fato de se tratar de um dicionário multilíngue coloca ambas as modelagens diante dos mesmos desafios pontuados na seção 2.1 deste trabalho. Para a modelagem do Turismo, a opção foi lançar mão dos *frames* cenários, utilizados pela FrameNet para organizar *frames* complexos e/ou produtivos, ou seja, aqueles que se relacionam com outros domínios e/ou que podem ser subdivididos em vários outros *frames*. Sua principal função é atuar na sistematização dos *frames* – servindo para orientar tanto usuários humanos quanto máquinas, daí o fato de, em vários casos, serem não lexicais. Dessa maneira, a organização dos *frames* do domínio do futebol no Kicktionary poderia acontecer através de *frames* cenários.

Outro ponto a ser destacado é a escolha pelos conceitos *frame* e *synset*. Como já discutido, ainda que tanto a FrameNet – representando aqui a teoria da Semântica de *Frames* – quanto a WordNet sejam redes semântico-lexicais, a abordagem adotada por cada uma é diferente. O fato de especificar as variadas relações lexicais de um mesmo item e também por organizar o vocabulário por classes de palavras faz com que a WordNet se destaque diante da base de dados

⁴⁰ “There is a common nonlinguistic use of this word where it refers not to what a piece of language evokes in the mind of an interpreter, but what kinds of conceptual structures an interpreter invokes to make sense of some experience. This nonlinguistic use of the concept has a long history (Bartlett, 1932; Piaget, 1971; Mill, 1846; see Fillmore, 1985), but it flourished in the cognitive and social sciences in the 1970s. The idea behind this work is that people understand what they are observing by calling on memories of past experiences, or by construing what is observed as instances of or variations from structures of belief and experience that could be used to make sense of what they have heard or observed. These structures in the mind are called frames, and these are not necessarily connected in specific ways with language”.

da FrameNet se o foco de análise for o fazer lexicográfico. Por mais que os dados estejam anotados, a FrameNet, até então, não sistematiza as relações lexicais e não apresenta, com o mesmo rigor⁴¹, as ULs por classes de palavras lexicais.

Nesse sentido, Schmidt localizou, na teoria da WordNet, uma possibilidade de efetivar essas questões esperadas por um dicionário. Como era preciso um modo de alinhar os dados das três línguas – alemão, francês, inglês –, o autor adapta o conceito de *synset* da WordNet. Não somente o conjunto de sinônimos de uma mesma língua mas também os correspondentes das outras línguas estariam incluídos num mesmo *synset*. O *frame* no Kicktionary assume um papel secundário. A função dele é mostrar ao usuário como o conhecimento daquele domínio se organiza com a visualização por cenas e *frames*.

O fato de o *synset* representar um conjunto de sinônimos organizados por classes de palavras é importante para o objetivo de um dicionário, mas, ao relacionar esses conjuntos de sinônimos para uma tarefa multilíngue, as particularidades que as línguas demonstram na lexicalização de conceitos não é demonstrada, e isso o *frame* poderia fazer. Por esse motivo, uma alternativa que exploraria melhor cada uma dessas duas teorias seria que as relações intralinguísticas fossem representadas pelos *synsets* enquanto as relações interlinguísticas se dessem via *frame* (SALOMÃO ET AL., 2011).

Em todos os trabalhos publicados, Schmidt destaca a importância de *corpus* para viabilizar as análises realizadas no Kicktionary. Porém, ao alinhar os dados em seu modelo especial de *synset*, que inclui os equivalentes entre línguas, o autor não especifica como o estudo de *corpora* contribuiu para a escolha dos melhores correspondentes. Se a análise semântica que Schmidt realiza nos *corpora* compilados para todas as línguas do dicionário se somasse à análise sintática, tal como é o procedimento de anotação da FrameNet, as evidências sintáticas especificadas nos padrões de valência poderiam contribuir para as informações trazidas pela valência semântica.

Essa questão pode ser exemplificada com os *frames* *Turismo_por_turista* e *Atração_turística*, duas perspectivas distintas do

⁴¹ A WordNet organiza os *synsets* por classes de palavras lexicais, a FrameNet tem como objetivo exibir as palavras que evocam *frames*, nesse sentido, independentemente da classe gramatical assumida, todas as palavras apresentam o mesmo *status* de predicadores capazes de evocar *frames*. Porém, ao realizar a busca, são assinaladas as Unidades Lexicais por classes de palavras, daí dizer que, ainda que essa questão seja pontuada, o tratamento não é feito com o mesmo rigor que o realizado pela WordNet.

frame Turismo_de_atração. Ainda que ambos os *frames* tenham os EFs Turista e Atração como nucleares, a perspectiva assumida para cada um não é a mesma, e isso fica evidenciado pela anotação sintática realizada pela FN. Em Turismo_por_turista, o EF Turista atua com função de Externo, e o EF Atração, como Objeto Direto, já, no *frame* Atração_turística, o EF Turista age como Objeto Direto ou Indireto, e a Atração, como Externo. Assim, caso fossem anotados apenas os papéis semânticos, eles indicariam que essa combinação de EFs se dirige a um mesmo *frame*, o que não acontece.

Como o autor destaca que contribuições dessa natureza são bem-vindas para o aperfeiçoamento desse recurso lexical, essas iniciativas valorizariam a potencialidade que a FrameNet e WordNet apresentam quando atuam em conjunto.

4.3 O papel da WordNet no copa 2014

Modelar o cenário do turismo para o dicionário Copa 2014 confirmou a importância de entender e, ao mesmo tempo, sistematizar o comportamento das palavras ao estruturar a experiência turística. Muitas vezes, um lexema assumiu significados específicos, atuando como diferentes Unidades Lexicais. Palavras como *check in* e *reservar*, por exemplo, participam de diferentes situações. Chegar ao aeroporto é diferente de chegar ao local da atração. A mesma situação para o lexema *reservar*. Reservar um hotel não é a mesma coisa que reservar um passeio turístico. Essas distinções semânticas na valência desses predicadores pontuam a atuação em eventos diferentes, que, por isso, requerem complementos específicos. A existência desse fenômeno é enfatizada no trabalho com *frames*, e isso não fica marcado no Kicktionary com o uso de *synsets* no tratamento das relações interlinguísticas. O *synset* especifica um sentido ao agrupar as palavras, e o usuário não tem a dimensão das outras possibilidades, pois essa marcação é perdida ao alinhar as línguas por *synsets*.

No dicionário Copa 2014, um papel importante que a WordNet pode ocupar é o de garantir as relações intralinguísticas. Hoje, a FrameNet não tem como propósito agrupar as Unidades Lexicais por classes de palavras lexicais nem sistematizar relações, como sinonímia e antonímia. Unidades Lexicais que explicitam

essa questão são as relacionadas ao lema *check in* no banco de dados da língua portuguesa. Enquanto, no português, o lema *check in* só adota comportamento verbal em construções com verbo suporte, no inglês, pode ser utilizado como verbo ou como nome. Dessa forma, caso as classes de palavras não sejam levadas em consideração de maneira sistematizada quando da constituição do recurso Copa 2014, como *check in.n* evoca, em português, o *frame* de *Chegada_do_turista_alojamento* e, em inglês, tanto *check in.n* e *check in.v* fazem isso, o usuário poderia ser levado a pensar que tais palavras são cognatos perfeitos nas duas línguas. Adicionar a estruturação em classes de palavras, juntamente com a consideração da valência sintático-semântica das ULs, faria com que o dicionário apontasse ao usuário que os sinônimos do verbo *check in* do inglês são *registrar-se* ou *fazer check in*.

Para usufruir aquilo que cada uma dessas teorias oferecem como redes que tratam o léxico das línguas, o Copa 2014 pode estruturar *synsets* para especificar o comportamento lexical interno de cada língua do dicionário, e, para viabilizar o alinhamento entre os sistemas linguísticos, o ideal é utilizar *frames*, possibilitando que estejam mantidas as particularidades de cada sistema linguístico ao perspectivar os *frames* tanto do domínio do turismo quanto dos demais.

5 CONCLUSÕES

A investigação inicial na qual esta proposta esteve engajada alcançou formulações importantes, que se podem nomear como diretrizes para o desenvolvimento do dicionário Copa 2014. Por mais que este trabalho esteja afiliado aos estudos linguísticos cognitivistas, o contato com outros ramos de pesquisa, como a Lexicografia, a Linguística Computacional e o Turismo mostrou a necessidade de pesquisar sobre cada uma dessas ciências. Essa interdisciplinaridade requereu um cuidado maior, o de manter este trabalho dentro do enfoque escolhido sem negligenciar a complexidade e as especificidades de cada área.

Quanto à primeira pergunta – *em que medida os frames do domínio turístico modelados com corpora compilados da língua portuguesa do Brasil servem para representar os frames deste domínio para as demais línguas do dicionário?* – chegou-se à conclusão de que os *frames* do domínio do turismo estruturados com *corpora* de língua portuguesa podem ser utilizados como moldes para o desenvolvimento dos *frames* das demais línguas. Isso se deveu ao caráter transcultural da atividade turística, pois, ainda que as culturas envolvidas em cada sistema linguístico demonstrem maneiras específicas para conceptualizar conceitos e experiências, o evento do Turismo perpassa essas particularidades, fazendo com que essa atividade possa ser moldada em torno de um *frame* padrão. Desse modo, modelar os *frames* com os dados do português brasileiro foi suficiente para garantir a estruturação dos *frames* das demais línguas, ainda que, para cada empreendimento, seja necessária a compilação de *corpora* específicos bem como seu estudo e sua análise.

Essa resposta possibilitou traçar o caminho para refletir a respeito da viabilidade do *frame* como interlíngua, questionada pela segunda pergunta – *Como a FrameNet responde aos desafios colocados na estruturação de recursos lexicais multilíngues? É possível utilizar frames como interlíngua?* A FrameNet hoje não fornece subsídio para muitos dos desafios colocados para a estruturação de dicionários eletrônicos multilíngues, entretanto isso não impede explorar os *frames* já como interlíngua para a confecção do Copa 2014. As características especiais colocadas tanto para o domínio do Turismo quanto para o Futebol e a Copa do

Mundo eliminam, por diferentes motivos, as divergências socioculturais, o que faz com que os *frames* tenham o mesmo *status* de nuclearidade. Assim sendo, as diferenças podem ser apresentadas nos padrões de realização, que podem se distinguir quanto às valências semânticas e sintáticas, mas não quanto à modelagem do evento.

Dessa forma, a relação de correspondência entre *frames* criada pela FrameNet Brasil guiará o alinhamento entre os sistemas linguísticos. Acerca das possíveis diferenças de lexicalização e, conseqüentemente, a busca pelos equivalentes de tradução, a proposta é explorar os padrões de valência proporcionados pela análise em três camadas da FN Br na tentativa de que contribuam para escolher as traduções. Há de se ressaltar que aqui as traduções se aproximam mais de paráfrases do que de equivalentes estritos, propriamente ditos, dialogando com Boas (2009) e Dias da Silva et al (2007) que ressaltam o papel das paráfrases em recursos lexicais alinhados via interlíngua.

Quanto à terceira pergunta – *Que avaliação se pode fazer do Kicktionary como produto que utiliza frames e synsets para o tratamento lexicográfico multilíngue?* – o modelo de recurso lexical multilíngue proposto por Schmidt no Kicktionary aponta, sem desconsiderar o prestígio desse recurso, que uma revisão teórica poderia contribuir muito para aperfeiçoar esse dicionário temático. A junção de duas teorias que concebem a informação lexical de forma diferente pode ser mais eficaz se levado em conta o que cada teoria se empenha em fazer. Nesse sentido, a WordNet é eficaz no tocante à organização das relações lexicais bem como a separação em categorias de palavras lexicais, medidas necessárias para a criação de dicionários. Já a FrameNet, por priorizar as relações semânticas entre os itens lexicais e organizar o léxico em torno dos elementos que estruturam um dado *frame*, como os participantes e as circunstâncias de um dado evento, pode contribuir mais do que a função assumida no Kicktionary, que, de modo geral, utilizou os *frames* para organizar o domínio do futebol na interface do dicionário.

A potencialidade do *frame* poderia ser explorada no Kicktionary para garantir as relações interlinguísticas assim como será no Copa 2014. Desse modo, é interessante explorar como os recursos fornecidos pela FrameNet podem contribuir. O procedimento de anotação em três camadas da FN pode ajudar a encontrar os equivalentes de tradução, e, para isso, tanto a anotação semântica quanto a sintática são úteis. Além desse critério de análise, o potencial do *frame* como

interlíngua também é uma oportunidade a ser considerada, já que a modelagem do futebol é concebida da mesma forma pelas diferentes culturas.

Num segundo momento, a distinção *cena* e *frame* utilizada por Schmidt para organizar os eventos do futebol reacende uma dicotomia superada pelos teóricos fundadores da Linguística Cognitiva. Um dos pilares dessa agenda de pesquisa é a associação entre conhecimento linguístico e conhecimento enciclopédico. Por esse motivo, não cabe dentro desse programa utilizar distinções que separem dicionário de enciclopédia. Fillmore, ao fazer essa distinção, não pretendia criar essa dicotomia, mas possibilitar mostrar os aspectos predominantes entre os diferentes tipos de *frames*, como aqueles de cunho mais abstrato daqueles mais específicos. Entretanto, como ele mesmo pondera, esses conceitos assim como outros possibilitaram uma confusão terminológica (Cf. FILLMORE & ANDOR, 2010). A melhor alternativa, diante dessa discussão, cara aos estudos cognitivos da linguagem, é lançar mão de *frames* cenários para organizar o conhecimento do futebol, tal como realizado para estruturar os *frames* do domínio do Turismo, esse tipo de *frame* modela *frames* abstratos e específicos de um mesmo domínio.

Como visto, as respostas para essas perguntas são de utilidade não só para a estruturação do dicionário Copa 2014 como recurso eletrônico multilíngue mas também podem servir para repensar a lexicografia como um todo, inserindo novas técnicas no fazer lexicográfico, o que de alguma forma caminha no sentido de tornar palpáveis alguns dos “sonhos dos lexicógrafos na era do dicionário eletrônico”, conforme Schryver (2007) já avaliara anteriormente.

Além dos estudos da linguagem, de modo geral, e da lexicografia, de modo específico, este trabalho também pode contribuir para a pesquisa relacionada às atividades turísticas. Ainda que o intuito não tenha sido o de atingir o rigor necessário para as terminologias da área, o estudo em *corpora* fornece uma pesquisa fundada em informação empírica, o que colabora para pesquisas desse ramo. A realização deste trabalho reforçou a dificuldade de unidade terminológica apontada por especialistas da área. Por mais que seja um ramo de pesquisa recente, o Turismo é relevante para diferentes domínios. Como mostrado pela modelagem do *frame* `Cenário_do_turismo`, essa área estabelece forte contato com outras áreas. Seja por comércio, negócios, eventos, religião, estudo ou descanso e entretenimento, inúmeros são os motivos que levam as pessoas a se deslocarem e aproveitarem as atividades turísticas que um local tem a oferecer.

Dessa forma, este trabalho exhibe um novo panorama para os especialistas refletirem sobre a experiência turística.

O recorte realizado por esta dissertação é parte de um projeto maior que necessita tratar de forma sistemática outras questões, como aquelas relativas à interface do dicionário e as que dizem respeito às especificidades do dicionário, que Borba (2003) nomeia como questões de nível técnico, embora também estejam amparadas por suporte teórico. Assim, é importante salientar que esta dissertação de mestrado não teve o propósito de garantir vocabulário representativo para as entradas do dicionário, ainda que as Unidades Lexicais e as análises de sentenças atuem como exemplares para as entradas lexicais, e também não teve como escopo o aprofundamento de questões técnico-computacionais sobre construção de dicionários. Esse tratamento está sendo realizado por outros pesquisadores da equipe FrameNet Brasil e também pela equipe parceira, o projeto FrameCorp.

A contribuição possibilitada com a estruturação do `Cenário_do_turismo` segue em andamento. Outros *frames*, como o `Cenário_do_turismo_planejamento` e o `Turismo_de_evento`, estão sendo aprofundados por outros colegas, o que fará com que o `Cenário_do_turismo` seja aperfeiçoado e ampliado em trabalhos posteriores.

Ademais, os desafios colocados para que a rede semântica FrameNet possa contribuir de forma efetiva para a criação de dicionários eletrônicos multilíngues necessita de outras investigações. Para trabalhos futuros, ressalta-se a importância de revisar a teoria da Linguística Cognitiva na busca por suporte teórico que seja capaz de responder a esses diversos desafios. É de se ressaltar que essas questões não estão restritas à rede semântica FrameNet, mas se colocam como obstáculos que deverão ser respondidos para o avançar do conhecimento nos estudos cognitivos da linguagem para posteriormente serem processados por máquinas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ATKINS, S. B. T.; RUNDELL, M.; SATO, H. The contribution of Framenet to practical lexicography. In: **International Journal of Lexicography**, v. 16, n. 3, Oxford: University Press, 2003, p. 333-357;
- ATKINS, S.B.T. & RUNDELL, M. **The Oxford Guide to Practical Lexicography**. Oxford: University Press, 2008;
- BARRETO, M. **Manual de iniciação ao estudo do turismo**. Campinas: Papyrus, 1995.
- BERTOLDI, A. **Semântica de Frames e recursos lexicais jurídicos: um estudo contrastivo**. São Leopoldo: UNISINOS, 2011. 136 p. Tese (doutorado). Programa de Pós-graduação em Linguística Aplicada da Universidade do Vale dos Sinos, São Leopoldo, 2011;
- BICK, E. **The Parsing System PALAVRAS: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework**. Arhus, Arhus University, 2000;
- BORBA, F. S. **Organização de dicionários: uma introdução à lexicografia**. São Paulo: Ed. UNESP, 2003;
- BOAS, H. Semantic frames as interlingual representations for multilingual lexical databases. In: BOAS, H. C. (Ed.). **Multilingual FrameNets in computational lexicography: Methods and applications**. Berlin/New York: Mouton de Gruyter, 2009, p. 59-100;
- CHISHMAN, R. L. O.; BERTOLDI, A.; LERMEN, L.; PADILHA, J. G. Corpus e Anotação Semântica: um Experimento para a Língua Portuguesa a partir da Semântica de Frames. In: **WebMedia 2008 - XIV Simpósio Brasileiro de Sistemas Multimídia e Web - V Workshop de Tecnologia e Informação**, 2008, Vila Velha. Anais - Webmedia, v. II. 2008, p. 321-325;
- CHOMSKY, N. **Aspects of the Theory of Syntax**. Cambridge, Mass.: MIT. Press, 1965;
- COOPER, C. et al. **Turismo, princípios e práticas**. trad. Roberto Cataldo Costa. 2. ed. Porto Alegre: Bookmann, 2001;
- CRUSE, D. A. **Lexical semantics**. Cambridge, Cambridge University Press, 1986;
- DIAS-DA-SILVA, B. D. et al., Introdução ao processamento das línguas naturais e Algumas Aplicações. In: **Série de Relatórios do NILC**. NILC-TR-07-10. São Carlos-SP, Agosto, 2007, p. 58-77;
- DI-FELIPPO, A.; DIAS-DA-SILVA, B. C. A interlíngua da base lexical bilingue REBECA. In: **Revista de Estudos Linguísticos Veredas**, v.13, n.02. Juiz de Fora, Minas Gerais, Brasil. ISSN Online: 1982-2243 / ISSN Impressa: 1415-2533. 2009. p. 50-67;
- FAUCONNIER, G. **Mappings in Thought and Language**. Cambridge: Cambridge University Press, 1997;
- FELLBAUM, C. **WordNet: an electronic lexical database**. 2.ed. Cambridge Mass.: MIT Press. 1998;

_____. & MILLER, G. A. WordNet then and now. In: **Language Resources & Evaluation**. v. 41, 2007, p. 209-214;

FILLMORE, C.J. The case for case. In: BACH, E. & HARMS, E. (Ed.). **Universals in linguistic theory**. New York, Holt, Rinehart and Winston, 1968, p. 1-88;

_____. Scenes and frames semantics. In: ZAMPOLLI, A. (Ed.). **Linguistic Structures Processing: Fundamental Studies in Computer Science**, nº 59. Amsterdam: North Holland Publishing. 1977;

_____. Innocence: a second idealization for linguistics. **Proceedings of the Fifth Berkeley Linguistics Society**, 1979;

_____. Frame semantics. In: **Linguistics in the Morning Calm**. Seul: Hanshin Publishing Co., 1982, p.111-137;

_____. Frames and the semantics of understanding. In: **Quaderni di Semantica**. v.6, n.2, 1985, p. 222-254;

_____. Corpus linguistics or computer corpus linguistics. In: **Directions in corpus linguistics**. Proceedings of nobel symposium 82, Stockholm, Ed. Jan Svartvik, Berlim/Nova York, De Gruyter, 1992, p. 35-60.;

_____. ; JONHSON, C.; PETRUCK, M. Background to FrameNet. In: **International Journal of Lexicography**. Oxford University Press, v. 16 nº 3, 2003, p. 235-250;

_____.; PETRUCK, M. R. L.; RUPPENHOFER, J. & WRIGHT, A. FrameNet in action: the case of attaching. In: **International Journal of Lexicography**. Oxford University Press, v. 26, nº 3. 2003a;

_____.; CALLEJAS, C. M. B. Entrevista a Charles J. Fillmore, **Odisea**, nº 4, 2003b, p. 41-48. Disponível em: <http://www.ual.es/odisea/Odisea04_BretonesCalleja.pdf>. Acesso em: 10/10/2012.

_____. Frames Semantics. In: **Encyclopedia of Language and Linguistics**. 2.ed. Elsevier, 2006, p. 613-620;

_____.The Merging of Frames. In: FAVRETTI, R.R. (Ed.). **Frames, Corpora, and Knowledge Representation**, Bononia: University Press, 2008, p. 1-12;

_____. Border Conflicts: FrameNet Meets Construction Grammar. In: **EURALEX**, vº13 , Barcelona: Anais... Universitat Barcelona Fabra, 2008a;

_____. & ANDOR, J. Discussing frame semantics: The state of the art: An interview with Charles J. Fillmore. In: **Review of Cognitive Linguistics**, 8(1), 2010, p. 157 176;

GAMONAL, M. A., PERON, S.R., TORRENT, T. T. Equivalentes de Tradução entre Português, Espanhol e Inglês na Constituição de Dicionário Temático Trilíngue para a Copa 2014. In: **XI Encontro de Linguística de Corpus**, Universidade Federal de São Carlos. Anais do XI ELC 2012, 2012;

GOFFMAN, E. **Frame Analysis: An Essay on the Organization of Experience**. New York, NY et al.: Harper & Row, 1974;

GOLDBERG, A. **Constructions: A Construction Grammar Approach to Argument Structure**. Chicago: The University of Chicago Press, 1995;

_____. **Constructions at Work: The nature of generalization in language**. Oxford: Oxford University Press, 2006;

KILGARRIFF, A., RYCHLY, P., SMRZ, P., TUGWELL, D., The Sketch Engine. In: **Proceedings of Euralex**, Lorient, França. 2004, p. 105-116. Disponível em: <<http://www.sketchengine.co.uk/>>. Acesso em: 2 de junho de 2010;

LAKOFF, G. & JOHNSON, M. **Metaphors We Live By**. Chicago: The University of Chicago Press, 1980;

_____. **Women, Fire, and Dangerous Things: What Categories Reveal About the Mind**. Chicago: University of Chicago Press, 1987;

_____. & JOHNSON, M. **Philosophy in the Flesh**. New York: Basic Books. 1999;

LÖNNEKER-RODMAN, B. Multilinguality and FramNet. In: **Technical Report**. TR-07-001. Berkeley: ICSI, 2007;

MARRAFA, P. **WordNet do Português: uma base de dados de conhecimento linguístico**. Lisboa: Instituto Camões. 2001;

_____. et al., TemaNet – WordNets Temáticas do Português: Alimentação, Arte, Comunicação, Desporto, Educação, Geografia, Habitação, Relações Humanas, Saúde, Seres Vivos, Transportes e Turismo. In: **Instituto Camões**, 2006. Disponível em: <<http://www.instituto-camoes.pt/temanet>>. Acesso em: 30/08/12;

_____. et al., LexTec – Léxico Técnico do Português: Ambiente, Banca, Comércio, Construção Civil, Direito Comercial Internacional, Economia e Gestão de Empresas, Energia, Seguros, Turismo, Telecomunicações. In: **Instituto Camões**, 2009. Disponível em: <<http://instituto-camoes.pt/lextec>>. Acesso em: 30/08/12;

MILLER, G.A. et al. Introduction to WordNet: An On-line Lexical Database. **Five Papers on WordNet**. TM. CSL Report 43, Julho, 1990. Revisión de Marzo. 1993;

MILLER, G. A. **Wordnet: a Lexical Database for English**. Commun. ACM, 38(11):341, 1995;

MINSKY, M. **The Society of Mind**. London: Willian Heinemann Ltd, 1987;

PETRUCK, M. R. L. Frame Semantics. In: VERSCHUEREN, J. OSTMAN, J. & BLOMMAERT, J. (Eds.) **Handbook of Pragmatics**. Amsterdam & Philadelphia: John Benjamins, 1996;

_____. et al., Reframing FrameNet Data. In: WILLIAMS, G & VESSIER, S (Ed.). **Proceedings of The 11th EURALEX International Congress**, Lorient, France, 2004, p. 405-416;

REDDY, M. J. The conduit metaphor - A case of frame conflict in our language about language. In: Ortony, A. (Ed.), **Metaphor and Thought**. 1.ed. Cambridge: Cambridge University Press, 1979, p. 284-297;

RUPPENHOFER, J. et al., **FrameNet II: Extended theory and practice**. Berkeley: International Computer Science Institute, 2010. Disponível em: <<http://framenet.icsi.berkeley.edu/>>. Acesso em 25/08/11;

SALOMÃO, M.M.M. FrameNet Brasil: um trabalho em progresso. In: **Calidoscópio**, Vol. 7. nº 2, 2009, p. 171-182;

_____. Entrevista com Margarida Salomão. In: **Revista Investigações**, Vol. 23, nº 2, Julho, 2010. Disponível em: <<http://www.revistainvestigacoes.com.br>>. Acesso em 07/08/12;

- SALOMÃO, M. M. M.; TORRENT, T. T.; CAMPOS, F. C. A.; BRAGA, R. M. M. & VIEIRA, M. B. **Copa 2014 Framenet Brasil**. Projeto apresentado ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito do Edital Universal nº 14/2011. Juiz de Fora. 2011;
- SARDINHA, T. B. **Linguística de Corpus**. São Paulo: Manole. 2004;
- SATO, H. FrameSQL: A Software Tool for FrameNet. In: **Proceedings of ASIALEX 03**, Tokyo, Japan. Asian Association of Lexicography. 2003, p. 251–258;
- SATO, H. New Functions of FrameSQL for Multilingual FrameNets. In: **Proceedings of the Sixth International Language Resources and Evaluation Conference**. European Language Resources association, Marrakech, Morocco, 2008, p. 758-762;
- SCHMIDT, T. Interfacing Lexical and Ontological Information in a Multilingual Soccer FrameNet. In: **Proceedings of OntoLex 2006 – Interfacing Ontologies and Lexical Resources for Semantic Web Technologies**. Italia, 2006;
- _____. The Kicktionary: A Multilingual Resource of the Language of Football. In: REHM, G., WITT, A. & LEMNITZER, L. (Ed.). **Data Structures for Linguistic Resources and Applications**. Tübingen: Gunter Narr, 2007;
- _____. The Kicktionary: Combining Corpus Linguistics and Lexical Semantics for a Multilingual Football Dictionary. In: LAVRIC, E. et al. (Ed.). **The Linguistics of Football**. Tuebingen: Gunter Narr, 2008, p. 11–23;
- _____. The Kicktionary - a multilingual lexical resource of football language. In: BOAS, H. (Ed.). **Multilingual FrameNets - Methods and Applications**. Berlin/New York: Mouton de Gruyter, 2009, p. 101-132;
- SCHRYVER, G. M., Lexicographers' Dreams in the Electronic-Dictionary Age. In: **International Journal of Lexicography**. v.16, nº 2, 2003, p.143-198;
- SILVA, O. V. da; KEMP, S. R. A., A evolução histórica do turismo: da antiguidade clássica a revolução industrial – século XVIII. In: **Revista Científica Eletrônica de Turismo**. Ano V – Número 9 – Junho de 2008 – Periódicos Semestral, – ISSN: 1806-9169. 2009. p. 50-67;
- TALMY, L. Toward a cognitive semantics. V. I: **Concept structuring systems**. Cambridge, Mass.: MIT Press, 2000.
- TOSQUI, P. **Construção e Ancoragem Ontológica do Vocabulário Básico Bilingue do Turismo para Fins Didáticos**. Araraquara: UNESP, 2007. 246 p. Dissertação. Programa de Pós-graduação em Linguística e Língua Portuguesa da Faculdade de Letras da Universidade do Estado de São Paulo, Araraquara, 2007;
- USIEKIEWICZ, J. L., OLKO, M. **Multilingual Dictionaries on-line: reality and perspectives**. Warszawa, 2006;
- VOSEN, P. **EuroWordNet General Document**. Project LE2-4003 & LE4-8328 report, University of Amsterdam, 2002. Acesso em 25/05/12.

ANEXOS

Atração_em_lugar

[Lexical Unit Index](#)

Definição [Definition]:

Um **Lugar** possui ou abriga uma **Atração**. Geralmente, são recursos naturais ou artificiais, com valor histórico ou cultural, que promovem trânsito de pessoas para conhecer e explorar suas potencialidades.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction] Recursos naturais e/ou artificiais, que apresentam algum tipo de valor, como cultural, histórico ou ecológico.

Lugar [Place] Lugar que abriga ou possui uma **Atração**.

Semantic Type: Location
Periféricos [Non-Core]:

Descrição [Depictive] Características específicas da **Atração** ou do **Lugar**.

Duração [Duration] Período que especifica o tempo que o **Lugar** hospeda ou possui uma **Atração**.

Semantic Type: Duration

Finalidade [Purpose] Alguma ação específica que o **Lugar** pretende conquistar por abrigar ou possuir uma **Atração**.

Maneira [Manner] Maneira como o **Lugar** conquistou a posse ou passou a hospedar uma **Atração**.

Meio [Means] Ações ou meios que possibilitam ao **Lugar** abrigar ou possuir determinada **Atração**.

Razão [Reason] Motivo que faz com que o **Lugar** ou a **Atração** seja interessante para turistas.

Turista [Tourist] Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios que um **Lugar** possui.

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on: [Turismo de atração](#)

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Turismo de atração](#)

Has Subframe(s):

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Unidades Lexicais [Lexical Units]:

abrigar.v, haver.v, lugar.n, possuir.v, ter.v

Criado por [Created by] 99 on 12/10/2012 06:00:15 BRST Seg

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
abrigar.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 11:06:21 BRST Seg
haver.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 11:07:34 BRST Seg
lugar.n	Created	Entrada Lexical [Lexical entry]		103	02/06/2013 02:15:34 BRST Qua
possuir.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	99	12/10/2012 06:16:42 BRST Seg
ter.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 11:07:16 BRST Seg

Atração_turística

Definição [Definition]:

Um **Lugar** ou uma **Atração**, geralmente, apresenta recursos naturais ou artificiais com valor cultural ou histórico que atraem turistas de diferentes origens.
O Brasil OFERECE diferentes opções turísticas.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction] Geralmente, é um espaço físico que faz gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios, o que faz com que seja reconhecido socialmente por apresentar valor histórico ou cultural ou mesmo por oferecer recursos naturais ou artificiais. Em alguns momentos, os benefícios podem ser considerados os próprios atrativos para o **Turista**.
Belo Horizonte OFERECE muitos bares para turistas.

Lugar [Place]
 Semantic Type: Location
Turista [Tourist]
 Semantic Type: Sentient
 Periféricos [Non-Core]:

Lugar que abriga recurso natural ou artificial com valor cultural ou histórico.

Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios de uma **Atração** ou de um **Lugar**.

Acompanhante [Co-participant] Aquele que desfruta os benefícios de uma **Atração** juntamente com o **Turista**.
 Semantic Type: Sentient
Descrição [Deceptive] Características específicas da **Atração** ou do **Lugar**.

Duração [Duration] O período de duração da atividade turística.
 Semantic Type: Duration

Finalidade [Purpose] Alguma ação específica que o **Turista** pretende realizar com a atividade turística, entretenimento e descanso, por exemplo.

Frequência [Frequency] Frequência em que a **Atração** ou o **Lugar** é procurado por turistas ou quantidade de vezes que o **Turista** realiza determinada ação durante o período da estadia.

Maneira [Manner] Detalhes da atividade turística no sentido de especificar o modo como ocorreu ou o estado do **Turista**.
 Semantic Type: Manner
Meio [Means] Ações ou meios que possibilitam a atividade turística em determinada localidade.

Origem [Source] Ponto de partida do deslocamento do **Turista**. Geralmente, é o local onde o **Turista** reside e para onde retornará após a estadia no local turístico.
 Semantic Type: Locative_relation

Razão [Reason] Razão que faz com que o **Lugar** ou a **Atração** seja interessante para turistas.
 Semantic Type: State_of_affairs

Tempo [Time] Quando a atividade turística acontece.
 Semantic Type: Time

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Atração, Lugar}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
 Is Inherited by:
 Perspective on: [Turismo de atração](#)
 Is Perspectivized in:
 Uses:
 Is Used by:
 Subframe of: [Turismo de atração](#)
 Has Subframe(s):
 Precedes:
 Is Preceded by:
 Is Inchoative of:
 Is Causative of:
 See also:

Unidades Lexicais [Lexical Units]:

apresentar.v, atrair.v, destacar-se.v, oferecer.v, surpreender.v

Criado por [Created by] 103 on 01/28/2013 09:59:44 BRST Seg

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
apresentar.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 10:59:48 BRST Seg
atrair.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 10:45:33 BRST Seg
destacar-se.v	Created	Entrada Lexical [Lexical entry]		103	01/28/2013 11:03:59 BRST Seg
oferecer.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 10:56:08 BRST Seg

Cenário_do_turismo_chegada

[Lexical Unit Index](#)

Definição [Definition]:

Um **Turista** chega ao **Destino** e, posteriormente, ao local de hospedagem. Os meios de transporte utilizados podem ser vários, como automóvel e avião. A viagem pode acontecer com **Acompanhante** e conter bagagens.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:**Nuclear [Core]:**

Destino [Goal] Semantic Type: Goal	Localidade aonde o Turista chega. O ponto final do deslocamento.
Meio_de_transporte [Mode_of_transportation]	O meio de transporte utilizado pelo Turista para chegar até o Destino .
Turista [Tourist] Semantic Type: Sentient	Indivíduo ou grupo que chega à localidade onde acontecerá a atividade turística.

Periféricos [Non-Core]:

Acompanhante [Co-participant] Semantic Type: Sentient	Aquele que chega ao local onde acontecerá a atividade turística juntamente com o Turista .
Bagagem [Baggage]	O Turista , geralmente, leva para o Destino pertences, como roupas e outros acessórios.
Duração [Duration] Semantic Type: Duration	O período de duração da estadia no Lugar escolhido para fazer turismo.
Finalidade [Purpose]	Algo que o Turista pretende com a chegada ao local desejado.
Lugar [Place] Semantic Type: Locative_relation	Localização do Turista quando esta não for especificada pelo Destino .
Maneira [Manner] Semantic Type: Manner	Características que determinam o estado do Turista ao chegar ao local desejado ou detalhes que especificam como se deu a chegada.
Meio [Means]	Ações ou meios que possibilitam que o Turista chegue ao Lugar planejado para a atividade turística.
Origem [Source]	Localização de origem do Turista . Ponto inicial do deslocamento.
Razão [Reason]	Motivos que fazem com que o Turista chegue à região ou ao Destino planejado para o turismo.
Tempo [Time] Semantic Type: Time	Quando o Turista chega ao Lugar planejado para o turismo.
Trajectoria [Path] Semantic Type: Path	A rota na qual o deslocamento para o Destino acontece.
Velocidade [Speed] Semantic Type: Speed	A velocidade com a qual o Turista chegou ao Destino .

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
 Is Inherited by:
 Perspective on:
 Is Perspectivized in:
 Uses:
 Is Used by:
 Subframe of: [Cenário do turismo](#)
 Has Subframe(s): [Chegada do turista alojamento](#), [Chegada do turista localidade](#)
 Precedes: [Cenário do turismo estada](#)
 Is Preceded by: [Cenário do turismo planejamento](#)
 Is Inchoative of:
 Is Causative of:
 See also:

Unidades Lexicais [Lexical Units]:

chegar.v

Criado por [Created by] 103 on 02/06/2013 06:15:56 BRST Qua

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
chegar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 04:57:36 BRST Qui

Cenário_do_turismo_estada

[Lexical Unit Index](#)

Definição [Definition]:

Várias motivações fazem com que o **Turista** tenha uma experiência turística durante a estada em um **Lugar**: atrações, eventos, religião, negócios, estudo, aventura etc. Durante o período de estada, essas motivações são concretizadas. O **Turista** se interessa por outras atividades além daquelas que motivaram o deslocamento. Essas atividades, geralmente, projetam-se em atrações e/ou lugares reconhecidos pela humanidade por conter importância histórica ou cultural, recursos naturais e/ou artificiais.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction]

Geralmente, é um espaço físico que faz gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios, o que faz com que seja reconhecido socialmente por apresentar valor histórico ou cultural ou por oferecer recursos naturais ou artificiais.

Lugar [Place]

Semantic Type: Location

Espaço geográfico onde o **Turista** está localizado, pode tanto conter a **Atração** quanto ser o próprio atrativo.

Turista [Tourist]

Semantic Type: Sentient

Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios de uma **Atração** ou de um **Lugar** de modo geral.

Periféricos [Non-Core]:

Acompanhante [Co-participant]

Aquele que compartilha a experiência na estada com o **Turista**.

Duração [Duration]

Semantic Type: Duration

O período de duração da estada em determinada localidade.

Finalidade [Purpose]

Alguma ação específica que o **Turista** pretende realizar com a estada no **Lugar**, entretenimento, negócios, estudo etc.

Frequência [Frequency]

Frequência com que a **Atração** ou o **Lugar** são procurados por turistas ou quantidade de vezes em que o **Turista** realiza determinada ação durante o período turístico.

Maneira [Manner]

Semantic Type: Manner

Detalhes do período da estada no sentido de especificar o modo como se deu a atividade turística ou o estado do **Turista**.

Meio [Means]

Semantic Type: State_of_affairs

Ações ou meios que possibilitam a estada do **Turista** em determinada localidade.

Meio_de_transporte [Mode_of_transportation]

O meio de transporte utilizado pelo **Turista** para se deslocar durante o período de estada.

Origem [Source]

Semantic Type: Source

Ponto de partida do deslocamento do **Turista**. Geralmente, é o local onde o **Turista** reside e para onde retornará após a estada no local turístico.

Razão [Reason]

Semantic Type: State_of_affairs

Razão que faz com que o **Lugar** ou a **Atração** sejam interessantes para turistas.

Tempo [Time]

Semantic Type: Time

Quando a estada acontece.

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Atração, Lugar}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Cenário do turismo](#)

Has Subframe(s): [Turismo de atração](#), [Turismo de evento](#)

Precedes: [Cenário do turismo partida](#)

Is Preceded by: [Cenário do turismo chegada](#)

Is Inchoative of:

Is Causative of:

See also:

Unidades Lexicais [Lexical Units]:

estada.n, estadia.n, estar.v

Criado por [Created by] 103 on 02/05/2013 08:14:30 BRST Ter

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
estada.n	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 11:02:19 BRST Qui
estadia.n	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 11:03:37 BRST Qui
estar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 11:04:19 BRST Qui

Cenário_do_turismo_partida

Definição [Definition]:

O **Turista** sai da **Origem** com o objetivo de chegar ao seu **Destino**, que, geralmente, é o local onde reside. Para que isso aconteça, é necessário que ele finalize a estadia no alojamento escolhido.

Muitos turistas SAEM do Rio de Janeiro pela Linha Vermelha

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Meio_de_transporte
[Mode_of_transportation]

O meio de transporte utilizado pelo **Turista** para deixar a **Origem**.
Os ônibus PARTEM todos os dias pela manhã

Origem [Source]

Localização do **Turista** antes de ele se deslocar.

Semantic Type: Location

Turista [Tourist]

Indivíduo ou grupo que parte do local turístico.

Semantic Type: Sentient

Periféricos [Non-Core]:

Acompanhante [Co-participant]

Aquele que se retira da **Origem** juntamente com o **Turista**.

Semantic Type: Sentient

Destino [Goal]

Localidade aonde o **Turista** pretende chegar. O ponto final do deslocamento.

Semantic Type: Goal

Distância [Distance]

A distância do deslocamento até o **Destino**.

Finalidade [Purpose]

O propósito do **Turista** ao se deslocar da **Origem**.

Semantic Type: State_of_affairs

Frequência [Frequency]

Quantas vezes um **Turista** ou um **Meio_de_transporte** se retiram da **Origem**.

Lugar [Place]

Localização a partir da qual se dá a partida do **Turista** quando esta não for especificada pela **Origem**.

Semantic Type: Location

Maneira [Manner]

Detalhes da realização da partida no sentido de especificar o estado do **Turista** e/ou o modo como se deu a partida em si.

Semantic Type: Manner

Meio [Means]

Ações ou meios que possibilitam a partida do **Turista**.

Tempo [Time]

O tempo em que ocorre a partida do **Turista**.

Semantic Type: Time

Trajectoria [Path]

A rota na qual o deslocamento para acontece.

Semantic Type: Path

Velocidade [Speed]

A velocidade com a qual o **Turista** se desloca.

Semantic Type: Speed

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Cenário do turismo](#)

Has Subframe(s): [Partida do turista alojamento](#), [Partida do turista localidade](#)

Precedes:

Is Preceded by: [Cenário do turismo estada](#)

Is Inchoative of:

Is Causative of:

See also:

Unidades Lexicais [Lexical Units]:

deixar.v, partir.v

Criado por [Created by] 103 on 02/01/2013 06:33:50 BRST Sex

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
deixar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 07:59:23 BRST Qui
partir.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 07:58:41 BRST Qui

Cenário_do_turismo_planejamento

[Lexical Unit Index](#)

Definição [Definition]:

O **Turista** procura conhecer mais sobre a região onde acontecerá o turismo. Diferentes informações são pesquisadas, como atrações locais, questões de culinária, saúde, documentação, clima, deslocamento e segurança. Além disso, podem ser realizadas compras de passagem, reserva de hotéis ou mesmo pacotes de viagens, incluindo tanto deslocamento e alojamento quanto passeios turísticos desejados. Essas questões podem ser consideradas o planejamento da **Atividade_turística**.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atividade_turística [Tourist Activity] Programação realizada pelo **Turista** para organizar as atividades de entretenimento que serão realizadas durante o turismo bem como precauções, como documentos e previsão do tempo, por exemplo.

Turista [Tourist] Indivíduo ou grupo que planeja e irá realizar **Atividade_turística**.
Semantic Type: Sentient
Periféricos [Non-Core]:

Acompanhante [Co-participant] Aquela que irá desfrutar os benefícios de uma **Atividade_turística** juntamente com o **Turista**.
Semantic Type: Sentient

Destino [Goal] Localidade aonde o **Turista** pretende chegar. O ponto final do deslocamento.
Semantic Type: Goal

Duração [Duration] O período de duração do planejamento da **Atividade_turística**.
Semantic Type: Duration

Finalidade [Purpose] Alguma ação específica que o **Turista** pretende realizar com a **Atividade_turística**, entretenimento e descanso, por exemplo.
Semantic Type: State_of_affairs

Lugar [Place] Local onde o **Turista** planeja a **Atividade_turística**.

Maneira [Manner] Características que determinam o estado do **Turista** ao planejar a **Atividade_turística** ou detalhes que especificam como se deu o planejamento.
Semantic Type: Manner

Meio [Means] Ações ou meios que possibilitam o planejamento da **Atividade_turística**.

Origem [Source] Localização de origem do **Turista**. Ponto inicial do deslocamento.
Semantic Type: Source

Razão [Reason] Motivos que fazem com que o **Turista** opte por determinada **Atividade_turística**.

Tempo [Time] Quando o planejamento da **Atividade_turística** acontece.
Semantic Type: Time

Relações Frame-frame [Frame-frame Relations]:

Inherits from: [Cenário do turismo](#)

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of:

Has Subframe(s): [Planejamento do turista](#), [Serviço turístico comprar](#)

Precedes: [Cenário do turismo chegada](#)

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Cenário_do_turismo

[Lexical Unit Index](#)

Definição [Definition]:

Atividade essencialmente humana cujo ator principal é um **Turista** interessado em conhecer e/ou usufruir os benefícios de uma **Atração** ou um **Lugar** de modo geral. Diferentes razões promovem esse trânsito de pessoas, como entretenimento, negócios, evento e estudos.

O Brasil ATRAI turistas pela diversidade cultural e geográfica

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction]

Geralmente, é um espaço físico que faz gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios, o que faz com que seja reconhecido socialmente por apresentar valor histórico ou cultural ou mesmo por oferecer recursos naturais ou artificiais. Em alguns momentos, os benefícios podem ser considerados os próprios atrativos para o **Turista**.

Lugar [Place]

Espaço geográfico visitado pelo **Turista**, pode tanto conter a **Atração** quanto ser o próprio atrativo.

Turista [Tourist]

Semantic Type: Sentient

Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios de uma **Atração** ou um **Lugar** de modo geral.

Periféricos [Non-Core]:

Acompanhante [Co-participant]

Aquele que compartilha a experiência turística com o **Turista**.

Duração [Duration]

Semantic Type: Duration

O período de duração da atividade turística.

Finalidade [Purpose]

Semantic Type: State_of_affairs

Alguma ação específica que o **Turista** pretende realizar com a atividade turística, entretenimento e descanso, por exemplo.

Localização_normal [Normal_location]

Semantic Type: Manner

Localização de origem do **Turista**. Ponto inicial do deslocamento.

Maneira [Manner]

Semantic Type: Manner

Características que determinam o estado do **Turista** ao ter experiências turísticas ou detalhes que especificam como se deu a atividade.

Meio [Means]

Ações ou meios que possibilitam a atividade turística em determinada localidade.

Razão [Reason]

Motivos que fazem com que a **Atração** ou o **Lugar** sejam interessantes para turistas.

Tempo [Time]

Semantic Type: Time

Quando o turismo acontece.

Tipo [Type]

Motivos específicos que deslocam o **Turista** para o **Lugar** ou **Atração**, que podem ser descritos como tipos de turismo.

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Atração, Lugar}

Relações Frame-frame [Frame-frame Relations]:

Inherits from: [Cenário_da_visita](#)

Is Inherited by: [Cenário_do_turismo_planejamento](#)

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of:

Has Subframe(s): [Cenário_do_turismo_chegada](#), [Cenário_do_turismo_estada](#), [Cenário_do_turismo_partida](#)

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Unidades Lexicais [Lexical Units]:

turismo.n

Criado por [Created by] 102 on 01/30/2013 11:54:06 BRST Qua

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
turismo.n	Created	Entrada Lexical [Lexical entry]		103	02/06/2013 05:18:41 BRST Qua

Chegada_do_turista_alojamento

Definição [Definition]:

O **Turista** chega ao **Alojamento** onde ficará hospedado por algum tempo.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Alojamento [Lodging] Semantic Type: Location	Hospedagem onde o Turista se aloja durante a estadia.
Meio_de_transporte [Mode_of_transportation] Semantic Type: Location	O meio de transporte utilizado pelo Turista para chegar ao Alojamento .
Turista [Tourist] Semantic Type: Sentient	Indivíduo ou grupo que chega ao Alojamento .

Periféricos [Non-Core]:

Acompanhante [Co-participant]	Aquele que chega ao local onde acontecerá a atividade turística juntamente com o Turista .
Finalidade [Purpose] Semantic Type: State_of_affairs	O propósito do Turista ao optar pelo Alojamento .
Lugar [Place] Semantic Type: Locative_relation	Localização do Turista quando esta não for especificada pelo Alojamento .
Maneira [Manner] Semantic Type: Manner	Detalhes da chegada ao Alojamento no sentido de especificar o estado do Turista e/ou o modo como aconteceu essa ação em si.
Meio [Means]	Ação ou meio que possibilita a chegada do Turista ao Alojamento .
Origem [Source] Semantic Type: Location	Localização do Turista antes de ele se deslocar.
Tempo [Time] Semantic Type: Time	O tempo em que ocorre a partida do Turista para o Alojamento .
Trajatória [Path] Semantic Type: Path	A rota na qual se dá o deslocamento para o Alojamento .
Velocidade [Speed] Semantic Type: Speed	A velocidade com a qual o Turista se desloca para o Alojamento .

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
Is Inherited by:
Perspective on:
Is Perspectivized in:
Uses:
Is Used by:
Subframe of: [Cenário do turismo chegada](#)
Has Subframe(s):
Precedes:
Is Preceded by: [Chegada do turista localidade](#)
Is Inchoative of:
Is Causative of:
See also:

Unidades Lexicais [Lexical Units]:

check-in.n, entrada.n, fazer check-in.v

Criado por [Created by] 103 on 02/07/2013 05:27:14 BRST Qui

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
check-in.n	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	02/14/2013 05:19:32 BRST Qui
entrada.n	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 05:40:23 BRST Qui
fazer check-in.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 05:43:15 BRST Qui

Chegada_do_turista_localidade

Definição [Definition]:

O **Turista** chega ao **Destino** onde realizará alguma atividade turística.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Destino [Goal] Semantic Type: Goal	Localidade aonde o Turista chega. O ponto final do deslocamento.
Meio_de_transporte [Means_of_transportation]	O meio de transporte utilizado pelo Turista para chegar ao Destino .
Turista [Tourist] Semantic Type: Sentient	Indivíduo ou grupo que chega à localidade onde acontecerá a atividade turística.

Periféricos [Non-Core]:

Acompanhante [Co-participant] Semantic Type: Sentient	Aquele que chega ao local onde acontecerá a atividade turística juntamente com o Turista .
Finalidade [Purpose] Semantic Type: State_of_affairs	Alguma ação específica que o Turista pretende realizar com a chegada à localidade.
Lugar [Place] Semantic Type: Locative_relation	Localização do Turista quando esta não for especificada pelo Destino .
Maneira [Manner] Semantic Type: Manner	Características que determinam o estado do Turista ao chegar ao Destino ou detalhes que especificam como aconteceu a chegada.
Meio [Means]	Ações ou meios que possibilitam que o Turista chegue ao Destino .
Origem [Source]	Localização de origem do Turista . Ponto inicial do deslocamento.
Quantidade [Quantity]	Especificação da quantidade de turistas.
Razão [Reason]	Motivos que fazem com que o Turista chegue à região ou ao Destino planejado para o turismo.
Tempo [Time] Semantic Type: Time	Quando o Turista chega ao Destino .
Trajectoria [Path] Semantic Type: Path	A rota pela qual o deslocamento para o Destino acontece.
Velocidade [Speed] Semantic Type: Speed	A velocidade com a qual o Turista chegou ao Destino .

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
Is Inherited by:
Perspective on:
Is Perspectivized in:
Uses:
Is Used by:
Subframe of: [Cenário do turismo chegada](#)
Has Subframe(s):
Precedes: [Chegada do turista alojamento](#)
Is Preceded by:
Is Inchoative of:
Is Causative of:
See also:

Unidades Lexicais [Lexical Units]:

aportar.v, desembarcar.v, desembarque.n

Criado por [Created by] 103 on 02/06/2013 06:41:22 BRST Qua

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
aportar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 05:12:10 BRST Qui
desembarcar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 05:11:27 BRST Qui
desembarque.n	Created	Entrada Lexical [Lexical entry]		103	02/08/2013 04:39:00 BRST Sex

Partida_do_turista_alojamento

Definição [Definition]:

O **Turista** finaliza sua hospedagem no local escolhido como **Alojamento**. Procedimentos específicos costumam ser necessários, como check out no hotel. No hotel, o **CHECK-OUT** deve ser realizado até meio dia.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Alojamento [Lodging] Semantic Type: Location	Local onde o Turista se aloja durante a estadia.
Meio_de_transporte [Mode_of_transportation] Semantic Type: Sentient	O meio de transporte utilizado pelo Turista para deslocar-se do Alojamento .
Turista [Tourist] Semantic Type: Sentient	Indivíduo ou grupo que sai do local do Alojamento e se dirige ao Destino .
Periféricos [Non-Core]:	
Acompanhante [Co-participant] Semantic Type: Sentient	Aquele que se retira do Alojamento juntamente com o Turista .
Bagagem [Baggage]	O Turista , geralmente, sai do Alojamento com pertences, como roupas e outros acessórios.
Destino [Goal]	Localidade aonde o Turista pretende chegar. O ponto final do deslocamento.
Finalidade [Purpose] Semantic Type: State_of_affairs	O propósito do Turista ao se deslocar do Alojamento .
Lugar [Place] Semantic Type: Locative_relation	Localização a partir da qual se dá a partida do Turista quando esta não for especificada pelo Alojamento .
Maneira [Manner] Semantic Type: Manner	Detalhes da realização da saída do Alojamento no sentido de especificar o estado do Turista e/ou o modo como aconteceu essa ação em si.
Meios [Means]	Ação ou meios que possibilitam a partida do Turista do Alojamento .
Tempo [Time] Semantic Type: Time	O tempo em que ocorre a partida do Turista do Alojamento .
Trajectoria [Path] Semantic Type: Path	A rota pela qual o deslocamento para o local desejado acontece.
Velocidade [Speed] Semantic Type: Speed	A velocidade com a qual o Turista se desloca do Alojamento .

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
Is Inherited by:
Perspective on:
Is Perspectivized in:
Uses:
Is Used by:
Subframe of: [Cenário do turismo partida](#)
Has Subframe(s):
Precedes: [Partida do turista localidade](#)
Is Preceded by:
Is Inchoative of:
Is Causative of:
See also:

Unidades Lexicais [Lexical Units]:

check-out.n, saída.n

Criado por [Created by] 103 on 02/05/2013 04:42:14 BRST Ter

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
check-out.n	Created	Entrada Lexical [Lexical entry]		103	02/05/2013 05:21:24 BRST Ter
saída.n	Created	Entrada Lexical [Lexical entry]		103	02/05/2013 05:15:20 BRST Ter

Partida_do_turista_localidade

Definição [Definition]:

Tendo finalizado o período de estadia no local turístico, o **Turista** parte rumo ao seu **Destino**.
Hoje pela manhã, **Vinícius** **EMBARCOU** com **Thiago e Carolina** para **Natal**.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Meio_de_transporte [Mode_of_transportation] O meio de transporte utilizado pelo **Turista** para chegar até o **Destino**.

Origem [Source] Localização do **Turista** antes de ele se deslocar.

Turista [Tourist] Indivíduo ou grupo que parte do local turístico para seu **Destino**.

Semantic Type: Sentient

Periféricos [Non-Core]:

Acompanhante [Co-participant] Aquela que se retira da **Origem** juntamente com o **Turista**.

Semantic Type: Sentient

Destino [Goal] Localidade aonde o **Turista** pretende chegar. Geralmente, será seu ponto inicial antes de ter realizado o deslocamento.

Semantic Type: Goal

Distância [Distance] A distância do deslocamento a partir da **Origem**.

Finalidade [Purpose] O propósito do **Turista** ao se deslocar da **Origem**.

Semantic Type: State_of_affairs

Lugar [Place] Localização a partir da qual se dá a partida do **Turista** quando esta não for especificada pela **Origem**.

Semantic Type: Locative_relation

Maneira [Manner] Detalhes da realização da partida da **Origem** no sentido de especificar o estado do **Turista** e/ou o modo como se deu esse deslocamento em si.

Semantic Type: Manner

Meio [Means] Ação ou meios que possibilitam a partida do **Turista** para o **Destino**.

Tempo [Time] O tempo em que ocorre a partida do **Turista**.

Semantic Type: Time

Trajectoria [Path] A rota pela qual o deslocamento a partir da **Origem** acontece.

Semantic Type: Path

Velocidade [Speed] A velocidade com a qual o **Turista** se desloca a partir da **Origem**.

Semantic Type: Speed

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Meio_de_transporte, Turista}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Cenário do turismo partida](#)

Has Subframe(s):

Precedes:

Is Preceded by: [Partida do turista alojamento](#)

Is Inchoative of:

Is Causative of:

See also:

Unidades Lexicais [Lexical Units]:

check-in.n, embarcar.v, embarque.n

Criado por [Created by] 103 on 02/05/2013 06:07:27 BRST Ter

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
check-in.n	Created	Entrada Lexical [Lexical entry]		103	02/05/2013 07:45:32 BRST Ter
embarcar.v	Created	Entrada Lexical [Lexical entry]		103	02/07/2013 09:08:58 BRST Qui
embarque.n	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	02/07/2013 09:00:33 BRST Qui

Planejamento_do_turista

Definição [Definition]:

O **Turista** pesquisa sobre o **Destino** onde acontecerá a **Atividade_turística**.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atividade_turística [Tourist Activity] Programação realizada pelo **Turista** para organizar as atividades de entretenimento que serão realizadas durante o turismo bem como precauções a serem tomadas, como documentos e compra de passagens.

Turista [Tourist] Indivíduo ou grupo que procura informações para preparar a **Atividade_turística**.

Semantic Type: Sentient
Periféricos [Non-Core]:

Acompanhante [Co-participant] Aquele que planeja a **Atividade_turística** juntamente com o **Turista**.

Semantic Type: Sentient

Destino [Goal] Localidade aonde o **Turista** pretende chegar. O ponto final do deslocamento.

Semantic Type: Goal

Duração [Duration] O período de duração do planejamento da **Atividade_turística**.

Semantic Type: Duration

Finalidade [Purpose] Alguma ação específica que o **Turista** pretende realizar com a **Atividade_turística**, entretenimento e descanso, por exemplo.

Semantic Type: State_of_affairs

Lugar [Place] Local onde o **Turista** planeja a **Atividade_turística**.

Semantic Type: Locative_relation

Maneira [Manner] Características que determinam o Estado do **Turista** ao planejar a **Atividade_turística** ou detalhes que especificam como se deu o planejamento.

Semantic Type: Manner

Meio [Means] Ações ou meios que possibilitam o planejamento da **Atividade_turística**.

Origem [Source] Localização de origem do **Turista**. Ponto inicial do deslocamento.

Semantic Type: Source

Razão [Reason] Motivos que fazem com que o **Turista** opte por determinada **Atividade_turística**.

Tempo [Time] Quando o planejamento da **Atividade_turística** acontece.

Semantic Type: Time

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Cenário do turismo planejamento](#)

Has Subframe(s):

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Serviço_turístico_comprar

Definição [Definition]:

O **Turista** realiza as compras necessárias para o planejamento da Atividade_turística.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Serviço_turístico [Tourist_service] Serviços ou produtos adquiridos pelo **Turista** quando do planejamento de sua viagem ou durante a atividade turística.

Turista [Tourist] Indivíduo ou grupo que efetua compras de **Serviço_turístico**.

Semantic Type: Sentient

Periféricos [Non-Core]:

Acompanhante [Co-participant] Aquela que irá comprar o **Serviço_turístico** juntamente com o **Turista**.

Semantic Type: Sentient

Agência_de_turismo [Tourist_agency] Indivíduo ou organização que vende o **Serviço_turístico** para o **Turista**.

Semantic Type: Locative_relation

Destino [Goal] Localidade aonde o **Turista** pretende chegar. O ponto final do deslocamento.

Semantic Type: Goal

Dinheiro [Money] Aquilo que o **Turista** paga em troca do **Serviço_turístico**.

Duração [Duration] O período de duração do compra do **Serviço_turístico**.

Semantic Type: Duration

Finalidade [Purpose] Alguma ação específica que o **Turista** pretende realizar com a compra de determinado **Serviço_turístico**.

Semantic Type: State_of_affairs

Lugar [Place] Local onde o **Turista** compra o **Serviço_turístico**.

Semantic Type: Locative_relation

Maneira [Manner] Características que determinam o estado do **Turista** ao efetuar a compra do **Serviço_turístico** ou detalhes que especificam como se deu esse processo.

Semantic Type: Manner

Meio [Means] Ações ou meios que possibilitam a compra do **Serviço_turístico**.

Origem [Source] Localização de origem do **Turista**. Ponto inicial do deslocamento.

Semantic Type: Source

Razão [Reason] Motivos que fazem com que o **Turista** opte pela compra de determinado **Serviço_turístico**.

Tempo [Time] Quando a compra do **Serviço_turístico** acontece.

Semantic Type: Time

Relações Frame-frame [Frame-frame Relations]:

Inherits from:

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Cenário do turismo planejamento](#)

Has Subframe(s):

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Turismo_de_atração

[Lexical Unit Index](#)

Definição [Definition]:

A busca pelo entretenimento é, basicamente, o que desloca o **Turista** para um **Lugar**. Ele deseja conhecer e/ou usufruir os benefícios que um **Lugar** ou uma **Atração** podem oferecer. Os tipos de atrações podem ser descritos separadamente para enfatizar suas especificidades, como turismo balneário, que explora tanto recursos naturais, como praias, quanto recursos artificiais, como parques aquáticos. Ainda que sejam descritos em eventos separados, todos os tipos de atrações podem ser encontrados a partir de três perspectivas: Turismo_por_turista, Atração_turística e Atração_em_lugar.

As **ATRAÇÕES** turísticas **de Minas Gerais** encantam pelas paisagens históricas.

Tipo Semântico [Semantic Type]: Non-perspectivalized_frame

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction]

Semantic Type: State_of_affairs

Geralmente, é um espaço físico que faz gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios, o que faz com que seja reconhecido socialmente por apresentar valor histórico ou cultural ou por oferecer recursos naturais ou artificiais. Em alguns momentos, os benefícios podem ser considerados os próprios atrativos para o **Turista**.

O Rio de Janeiro SURPREENDE pelas belas praias.

Turista [Tourist]

Semantic Type: Sentient

Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios de uma **Atração** ou de um **Lugar** de modo geral.

Periféricos [Non-Core]:

Acompanhante [Co-participant]

Semantic Type: Sentient

Aquele que desfruta os benefícios de uma **Atração** ou de um **Lugar** juntamente com o **Turista**.

Duração [Duration]

Semantic Type: Duration

O período de duração de determinada atividade turística.

Finalidade [Purpose]

Semantic Type: State_of_affairs

Alguma ação específica que o **Turista** pretende realizar com a atividade turística, entretenimento e descanso, por exemplo.

Frequência [Frequency]

Quantidade de vezes em que o **Turista** realiza determinada ação durante o período da estada.

Lugar [Place]

Semantic Type: Location

Espaço geográfico visitado pelo **Turista**, pode tanto conter a **Atração** quanto ser o próprio atrativo.

Maneira [Manner]

Semantic Type: Manner

Detalhes da atividade turística no sentido de especificar o modo como ocorreu ou o estado do **Turista**.

Meio [Means]

Uma ação do **Turista** que possibilita a sua estada em determinada localidade.

Origem [Source]

Semantic Type: Source

Ponto de partida do deslocamento do **Turista**. Geralmente, é o local onde o **Turista** reside e para onde retornará após a estada no local turístico.

Quantidade [Quantity]

Especificação da quantidades de atrações.

Razão [Reason]

Semantic Type: State_of_affairs

Razão que faz com que o **Lugar** ou **Atração** sejam interessantes para turistas.

Tempo [Time]

Semantic Type: Time

Quando a atividade turística acontece.

Tipo [Type]

Especificação das atrações para mostrar sua importância para a sociedade e as características que as particularizam.

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Atração, Lugar}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
 Is Inherited by:
 Perspective on:
 Is Perspectivized in: [Atração em lugar](#), [Atração turística](#), [Turismo por turista](#)
 Uses:
 Is Used by:
 Subframe of: [Cenário do turismo estada](#)
 Has Subframe(s): [Atração em lugar](#), [Atração turística](#)
 Precedes:
 Is Preceded by:
 Is Inchoative of:
 Is Causative of:
 See also:

Unidades Lexicais [Lexical Units]:

atração.n

Criado por [Created by] 103 on 02/06/2013 11:09:23 BRST Qua

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
atração.n	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	02/06/2013 11:41:38 BRST Qua

Turismo_de_evento

[Lexical Unit Index](#)

Definição [Definition]:

Eventos que geram trânsito de pessoas de diferentes localidades. Ainda que as atividades turísticas não sejam o motivo principal do deslocamento, os visitantes, geralmente, buscam formas de entretenimento, como atrações turísticas nos locais que sediam os eventos.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Evento [Event] Semantic Type: Event	Evento do qual o Turista participa.
Turista [Tourist] Semantic Type: Sientient	Indivíduo ou grupo que participa de um Evento em algum Lugar e, nessa ocasião, fará atividades turísticas.

Periféricos [Non-Core]:

Acompanhante [Co-participant] Semantic Type: Sientient	Aquele que participa do Evento junto com o Turista .
Atração [Attraction]	Geralmente, é um espaço físico de interesse coletivo capaz de gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios. Em alguns momentos, os benefícios podem ser considerados os próprios atrativos para o Turista .
Duração [Duration]	O período de duração de determinado Evento .
Finalidade [Purpose]	Alguma ação específica que o Turista pretende realizar com a participação no Evento .
Frequência [Frequency]	Frequência com que o Evento acontece.
Lugar [Place] Semantic Type: Locative_relation	Lugar que está sediando um Evento .
Maneira [Manner] Semantic Type: Manner	Detalhes que especificam o modo como se deu o Evento ou o estado do Turista .
Razão [Reason]	Razão que faz com que o Evento promova trânsito de pessoas de diferentes origens.
Tempo [Time] Semantic Type: Time	Quando o Evento acontece.

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
 Is Inherited by:
 Perspective on:
 Is Perspectivized in:
 Uses:
 Is Used by:
 Subframe of: [Cenário do turismo estada](#)
 Has Subframe(s):
 Precedes:
 Is Preceded by:
 Is Inchoative of:
 Is Causative of:
 See also:

Turismo_por_turista

Definição [Definition]:

Um **Turista** visita ou experiencia o contato com uma **Atração**, reconhecida socialmente, geralmente, por apresentar valor histórico ou cultural e/ou oferecer recursos naturais ou artificiais.

Durante a Copa, os torcedores CONHECERAO o Pelourinho em Salvador.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Atração [Attraction]

Geralmente, é um espaço físico que faz gerar trânsito de pessoas de diferentes origens para conhecer ou usufruir seus benefícios, o que faz com que seja reconhecido socialmente por apresentar valor histórico ou cultural ou por oferecer recursos naturais ou artificiais. Em alguns momentos, os benefícios podem ser considerados os próprios atrativos para o **Turista**.

iris VISITOU as diversas praias de Florianópolis.

Lugar [Place]

Semantic Type: Location

Lugar que abriga recurso natural ou artificial com valor cultural ou histórico, que pode atuar como **Atração**.

Turista [Tourist]

Semantic Type: Sentient

Indivíduo ou grupo interessado em conhecer ou desfrutar os benefícios de uma **Atração** ou de um **Lugar** específico.

Periféricos [Non-Core]:

Acompanhante [Co-participant]

Semantic Type: Sentient

Aquele que desfruta os benefícios de uma **Atração** juntamente com o **Turista**.

Duração [Duration]

Semantic Type: Duration

O período de duração da atividade turística

Finalidade [Purpose]

Semantic Type: State_of_affairs

Alguma ação específica que o **Turista** pretende realizar com a atividade turística, entretenimento e descanso, por exemplo.

Frequência [Frequency]

Frequência com que a **Atração** ou o **Lugar** são procurados por turistas ou quantidade de vezes que o **Turista** realiza determinada ação durante o período da estada.

Maneira [Manner]

Semantic Type: Manner

Detalhes da atividade turística no sentido de especificar o modo como ocorreu ou o estado do **Turista**.

Meio [Means]

Ações ou meios que possibilitam a atividade turística em determinada localidade.

Meio de transporte

[Mode_of_transportation]

O meio de transporte utilizado pelo **Turista** para se deslocar durante as atividades turísticas.

Origem [Source]

Ponto de partida do deslocamento do **Turista**. Geralmente, é o local onde o **Turista** reside e para onde retornará após a estadia no local turístico.

Razão [Reason]

Semantic Type: State_of_affairs

Razão que faz com que o **Lugar** ou **Atração** sejam interessantes para turistas.

Tempo [Time]

Semantic Type: Time

Quando a atividade turística acontece.

Conjunto(s) de EF Nucleares [FE Core set(s)]:

{Atração, Lugar}

Relações Frame-frame [Frame-frame Relations]:

Inherits from:
 Is Inherited by:
 Perspective on: [Turismo de atração](#)
 Is Perspectivized in:
 Uses:
 Is Used by:
 Subframe of:
 Has Subframe(s):
 Precedes:
 Is Preceded by:
 Is Inchoative of:
 Is Causative of:
 See also:

Unidades Lexicais [Lexical Units]:

apreciar.v, conhecer.v, desfrutar.v, fazer tour.n, turista.n, visitante.n, visitar.v

Criado por [Created by] 103 on 01/28/2013 11:33:21 BRST Seg

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]	Rel Anotação [Annotation Report]	ID Anotador [Annotator ID]	Data Criação [Created Date]
apreciar.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 01:45:02 BRST Seg
conhecer.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 01:14:09 BRST Seg
desfrutar.v	Created	Entrada Lexical [Lexical entry]		103	02/01/2013 02:59:07 BRST Seg
fazer tour.n	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	02/08/2013 05:21:15 BRST Seg
turista.n	Created	Entrada Lexical [Lexical entry]		103	01/28/2013 01:47:59 BRST Seg
visitante.n	Created	Entrada Lexical [Lexical entry]		103	01/28/2013 02:07:06 BRST Seg
visitar.v	Created	Entrada Lexical [Lexical entry]	Anotação [Annotation]	103	01/28/2013 01:08:32 BRST Seg