

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

Ludmila Meireles Lage

**FRAMES E CONSTRUÇÕES: A IMPLEMENTAÇÃO DO CONSTRUCTICON
NA FRAMENET BRASIL**

Juiz de Fora
2013

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

**FRAMES E CONSTRUÇÕES: A IMPLEMENTAÇÃO DO CONSTRUCTICON
NA FRAMENET BRASIL**

Ludmila Meireles Lage

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

FRAMES E CONSTRUÇÕES: A IMPLEMENTAÇÃO DO CONSTRUCTICON NA FRAMENET BRASIL

Ludmila Meireles Lage

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

Profa. Dra. Heliana Ribeiro de Mello – UFMG

Profa. Dra. Thaís Fernandes Sampaio – UFJF

Juiz de Fora
Março de 2013

Enquanto um mapa representa a geografia e uma partitura representa uma melodia, a língua representa nossa interpretação da realidade.

Haiman

A meus mestres, meus pais

AGRADECIMENTOS

Início meus agradecimentos pelo ponto de partida dessa dissertação, a instituição que proporcionou essa, assim como tantas outras de minhas realizações, a Universidade Federal de Juiz de Fora, a qual, através do Programa de Pós-Graduação em Linguística, permitiu meu aprimoramento intelectual e apoiou-me nas múltiplas demandas que surgiam. Agradeço, igualmente, à Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior – CAPES – pelo suporte no decorrer da pós-graduação.

Ao longo do curso de mestrado, tive a oportunidade de estudar – na maioria dos casos, novamente – com professores que contribuíram em alto grau com meu enriquecimento teórico. Aos professores Cristina Name, Neusa Salim, Margarida Salomão, Sonia Bittencourt e Tiago Torrent, meu sincero agradecimento.

Agradeço à Margarida Salomão por apresentar-me o Projeto FrameNet nos idos de 2007. O tempo dedicado ao trabalho na Iniciação Científica foi essencial para a decisão acertada de seguir nesse projeto. Agradeço a generosidade em compartilhar seus conhecimentos e a capacidade de dar asas àqueles que te cercam.

A meu orientador, Tiago Torrent, agradeço por ter sido o responsável por me convocar para a pós-graduação. O trabalho de um pós-graduando, como fui diversas vezes lembrada, é árduo e “desesperadamente rápido”, mas poderia ter sido impossível, não fosse sua infinita disponibilidade, atenção, cuidado e estímulo. Agradeço imensamente pela parceria àquele que faz jus ao *status* de orientador.

No âmbito do Projeto FrameNet, não posso deixar de agradecer aos primeiros orientandos do projeto, Gabriela, Paula e Renata, junto às quais trabalhei como bolsista de Iniciação Científica. Em especial, agradeço aos primeiros colegas de Iniciação, Daniela, Maucha, Michele e Reinaldo, com os quais dividi as primeiras dúvidas e descobertas do universo dos *frames*. Por fim, agradeço aos colegas que me auxiliaram na composição dessa dissertação no domínio do Projeto *Frames* e Construções: Adrieli, Carina, Marcela e Natália, muito obrigada.

Cabe, em separado, agradecer à minha amiga Maucha por dividir comigo todas as experiências da pós-graduação. Obrigada por sua ajuda, pelas tardes e noites de trabalhos partilhadas, pelas reflexões e amizade. Espero sua companhia em mais uma longa jornada.

Finalmente, agradeço à minha admirável família, Maurício, Auxiliadora e Victor, a qual esteve sempre presente apoiando-me das maneiras mais diversas. A meu pai, agradeço por apoiar minhas decisões e acreditar em mim. À minha mãe, agradeço o aconchego, carinho e apoio incondicional. Vocês são os responsáveis por tornar a minha vida possível. Ao Victor, o meu amor, agradeço por ser o amigo de todas as horas e situações. Só foi possível superar as dificuldades, porque havia você ao fim do dia. Coube à sua presença amorosa tornar meus dias mais suaves e felizes.

Meireles Lage, Ludmila .

Frames e Construções: A implementação do Constructicon na
FrameNet Brasil / Ludmila Meireles Lage. -- 2013.

113 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. Gramática das Construções. 2. Semântica de Frames. 3.
FrameNet Brasil. 4. Constructicon. 5. Construções em Para (SN)
Infinitivo. I. Timponi Torrent, Tiago , orient. II. Título.

RESUMO

Esta dissertação propõe apresentar as discussões teórico-metodológicas que embasam a implementação, na FrameNet Brasil (FN-Br), de um repertório de construções disponível *online*: o Constructicon do português brasileiro. A iniciativa constitui-se em um subprojeto da FN-Br, a qual, por sua vez, visa à criação de um recurso lexical *online* para o mesmo idioma, baseado na Semântica de *Frames* e sustentado por evidência em *corpus*. Para implementar o Constructicon, onze construções da família *Para Infinitivo* foram escolhidas. A escolha se deve ao fato de que essas construções formam uma rede formal e funcionalmente motivada no português do Brasil, além de evocarem *frames* relacionados – tais construções compartilham o esquema sintático [SN V SA/SN para (SN) V_{inf}] e o *frame* de Finalidade ou algum *frame* a ele relacionado. Ao longo do trabalho, realizamos análises dessas construções e discutimos quais poderiam ser tratadas como um padrão de valência, isto é, ser anotadas utilizando-se da metodologia centrada em unidades lexicais da própria FrameNet, e quais precisariam ser descritas no Constructicon. Propõe-se, dessa forma, um diálogo entre a proposta goldbergiana para o tratamento das construções e aquela idealizada por Fillmore e colaboradores. Discute-se, ainda, de que maneira e em que medida construções de estrutura argumental e unidades lexicais verbais interagem no que tange aos *frames* que evocam. Por fim, esta dissertação estabelece critérios que devem ser aplicados às potenciais construções do português brasileiro de modo a definir se elas devem ser tratadas como padrões de valência ou se devem figurar no Constructicon.

Palavras-chave: 1. Gramática das Construções. 2. Semântica de Frames. 3. FrameNet Brasil. 4. Constructicon. 5. Construções em Para (SN) Infinitivo.

ABSTRACT

This dissertation presents the theoretical and methodological discussions supporting the implementation, in FrameNet Brasil (FN-Br), of a repertoire of constructions available online: the Brazilian Portuguese Constructicon. The initiative is a subproject of the FN-Br, which, in turn, aims at creating an online lexical resource for the same language, based on Frame Semantics and supported by *corpus* evidence. To implement the Constructicon, eleven constructions of the *Para* Infinitive Family were chosen. This choice is due to the fact that these constructions form a both formally and functionally motivated network in Brazilian Portuguese, besides evoking related frames – such constructions share the syntactic scheme [NP V AP/NP *para* (NP) V_{inf}] and the Purpose frame or some frame related to it. Along this work, we performed analyzes of these constructions and discussed whether they could be treated as a valence pattern, i.e. be annotated using the FrameNet methodology centered on lexical units; or be described in the Constructicon. It is proposed, therefore, a dialogue between the Goldberg proposal for the treatment of constructions and that designed by Fillmore et al. We also discuss how and to what extent argument structure constructions and verbal lexical units interact in terms of the frames they evoke. Finally, this dissertation establishes criteria which must be followed so as to determine whether a potential construction should be annotated as a valence pattern or accounted for in the Constructicon.

Keywords: 1. Construction Grammar. 2. Frame Semantics. 3. FrameNet Brasil. 4. Constructicon. 5. Para (NP) Infinitive Constructions.

SUMÁRIO

INTRODUÇÃO	12
1 METODOLOGIA	14
1.1 Enquadramento Metodológico	14
1.2 Corpus	16
1.2.1 <i>Constituição dos Corpora da FrameNet Brasil</i>	16
1.2.2 <i>Tratamento Computacional</i>	17
1.3 Métodos de Anotação na FrameNet Brasil	21
1.4.1 <i>Anotação Lexicográfica</i>	21
1.4.2 <i>Anotação de Texto Corrido</i>	23
1.4.3 <i>Anotação Construcional</i>	26
2 FRAMES E CONSTRUÇÕES	27
2.1 Semântica de Frames	27
2.1.1 <i>Origem e Principais Desdobramentos</i>	27
2.1.2 <i>FrameNet</i>	34
2.2 A Gramática das Construções	46
2.2.1 <i>A Gramática Cognitivista das Construções</i>	47
2.2.1.1 <i>Significados Verbais e Significados Construcionais</i>	48
2.2.1.2 <i>Motivação e Relações de Herança entre Construções</i>	52
2.2.2 <i>A Gramática das Construções Baseada em Unificação</i>	60
2.3 O Constructicon	65
3 AS CONSTRUÇÕES DA FAMÍLIA PARA INFINITIVO	74
3.1 Construções Tratáveis como Padrões de Valência	82
3.2 Construções Incluídas no Constructicon	85
4 IMPLICAÇÕES DO TRATAMENTO COMPUTACIONAL DAS	98
CONSTRUÇÕES DA FAMÍLIA PARA INFINITIVO	
4.1 O Princípio da Não-Redundância	98
4.2 Reflexões sobre O Conceito de Construção	102
4.3 Inconsistências Presentes no Modelo do Constructicon	105
5 CONCLUSÕES	108
REFERÊNCIAS BIBLIOGRÁFICAS	110

INTRODUÇÃO

A presente dissertação tem por objetivo apresentar as discussões teórico-metodológicas que embasam a implementação de um repertório de construções disponível *online*: o Constructicon do português brasileiro. Esse empreendimento constitui-se em um subprojeto da FrameNet Brasil, a qual também visa à criação de um recurso *online* para o mesmo idioma, esse, por sua vez, um repertório de unidades lexicais baseado na Semântica de *Frames* e sustentado por evidência em *corpus*.

Para realizar a implementação do Constructicon, onze construções da família *Para Infinitivo* foram escolhidas, quais sejam, i. Construção de Dativo com Infinitivo; ii. Construção Modal Epistêmica; iii. Construção Modal Deôntica; iv. Construção de Propósito Qualificado; v. Construção de *Gap* de Sujeito; vi. Construção de *Gap* de Objeto; vii. Construção de *Gap* de Adjunto; viii. Construção Volitiva; ix. Construção Manipulativa; x. Construção de Tempo Iminente; e xi. Construção de Aspecto Iterativo.

A escolha pela família *Para Infinitivo* decorre do fato de que ela configura uma rede formal e funcionalmente motivada no português do Brasil, além de evocar *frames* relacionados – suas construções compartilham o esquema sintático [SN V SA/SN para (SN) V_{inf}] e o *frame* de Finalidade, ou algum *frame* a ele relacionado. Afora isso, as construções dessa família apresentam tipologia construcional bastante diversificada no que tange aos graus de esquematicidade ou de especificidade lexical, o que favorece nossa tarefa de implantação do Constructicon, uma vez que essa grande variação nos permite avaliar diferentes tipos de construções e, assim, concluir a quais tipos cabe um tratamento construcional nos moldes da FrameNet.

Na dissertação, enquadrada como pesquisa do campo investigativo da Linguística Computacional bem como da Linguística de *Corpus*, apresentamos os procedimentos metodológicos seguidos pela FrameNet Brasil, desde a constituição dos *corpora*, passando pelo tratamento computacional por eles recebido, até os diferentes processos de anotação – incluindo o construcional, foco deste trabalho.

A discussão teórica que aqui será exposta traça o caminho percorrido a partir da inauguração da Linguística Cognitiva, ambiente em que surge a

Semântica de *Frames*, chegando às diversas versões das Gramáticas das Construções. Tratamos, especificamente, da Gramática Cognitivista das Construções, de Goldberg (1995, 2006) e da Gramática das Construções Baseada em Unificação, adotada por Fillmore. A vertente de Goldberg foi escolhida com base no fato de que foi empregada na descrição que Torrent (2009) realiza acerca da família *Para Infinitivo*. Por sua vez, a versão adotada por Fillmore é detalhada visto que nela se encontram as bases para o Constructicon de Berkeley.

Após serem descritas as onze construções da família *Para Infinitivo*, discute-se quais delas devem ser tratadas via Constructicon, e quais, por outro lado, se adéquam ao tratamento lexicográfico via padrões de valência.

De posse de tais análises, é preciso definir o que será considerado construção, no âmbito da FrameNet Brasil. Em outras palavras, quais os critérios serão empregados para se avaliar e definir os elementos que devem ser anotados no modo construcional. A discussão passa por questões como a Não-Redundância em banco de dados, além de apontar os principais problemas encontrados no Constructicon americano. Uma das razões em se estabelecerem os critérios de anotação via Constructicon parte da lacuna deixada pelos criadores do projeto de Berkeley, o qual não os estabelece de modo explícito.

1 METODOLOGIA

Este capítulo apresentará a metodologia aplicada nesta dissertação. Iniciaremos o capítulo expondo os principais pontos relativos à área investigativa em que esse trabalho se insere, ressaltando a abordagem adotada em nossas análises. Em seguida, apresentaremos o *corpus* empregado e o tratamento computacional por qual ele passa. Finalmente, descreveremos os métodos de anotação da FrameNet.

1.1 Enquadramento Metodológico

O presente trabalho pertence ao campo investigativo denominado Linguística Computacional, descrito por Othero e Menuzzi (2005) como o estudo encarregado do tratamento computacional da linguagem e das línguas naturais. Trata-se de uma área interdisciplinar a qual integra pesquisas em Linguística e em Computação, a fim de possibilitar “a construção de sistemas com capacidade de reconhecer e produzir informação apresentada em linguagem natural” (VIEIRA & LIMA, 2001, p.1).

Os primeiros estudos em Linguística Computacional são da década de 1950 e resultam de tentativas de se criar tradutores automáticos. Desde então, essa área do conhecimento vem se desenvolvendo cada vez mais devido ao crescimento dos estudos na área da Computação. À medida que a Informática avança, mais se exige e se espera da Linguística Computacional.

A Inteligência Artificial trabalha visando a facilitar a interação entre ser humano e máquina. Entre nós humanos, a maneira mais eficiente de comunicação ocorre por meio das línguas naturais. Nada mais coerente que transferir esse conhecimento à máquina, de modo a tornar tal interação mais amigável ao usuário. Entretanto, essa não é uma tarefa simples, uma vez que é preciso ensinar a uma máquina de raciocínio binário a processar uma língua real.

Para que o computador seja capaz de processar uma língua natural é preciso equipá-lo com um grande aparato de informações. Essas informações vêm das áreas tradicionais da Linguística, tais como a Fonética, a Fonologia, a Sintaxe e a Semântica. Conforme apontam Othero e Menuzzi,

todo esse conhecimento é utilizado para tentar processar (leia-se 'compreender e produzir') as línguas naturais e dominar o conhecimento linguístico de um ser humano (2005, p.22).

Duas subáreas de pesquisa se estruturam a partir da Linguística Computacional: a de Processamento de Linguagem Natural (PLN) e a de Linguística de *Corpus*. O campo do PLN visa à elaboração de programas que possam interpretar e gerar informações em linguagem natural (VIEIRA & LIMA, 2011). Seus estudos estão voltados à modelagem de *softwares* específicos, tais como tradutores automáticos, corretores ortográficos e gramaticais e geradores automáticos de resumos.

A Linguística de *Corpus*, por sua vez, é responsável pela compilação criteriosa de *corpora* contendo exemplares da língua pertencentes a diversos gêneros textuais. Os *corpora* compilados são utilizados em diversos tipos de pesquisa linguística. Este foi um dos campos da Linguística que mais se beneficiou do avanço da Informática, dado que o trabalho com *corpora* só foi possível através da tecnologia computacional.

A presente pesquisa se enquadra no âmbito da Linguística de *Corpus* e, como tal, realiza coleta e análise de *corpora* com auxílio de ferramentas eletrônicas. Nesse contexto, cabe uma reflexão acerca da aparente distinção entre análises quantitativas e qualitativas.

Ao traçar os procedimentos metodológicos de uma pesquisa, sempre deparamo-nos com a questão da abordagem adotada no trabalho, se quantitativa ou qualitativa. Contudo, as análises que apresentaremos ao longo dessa dissertação não podem ser classificadas como quantitativas ou qualitativas.

Schiffrin (1987, p.66-67) discute a validade de enquadramento em uma das duas abordagens considerando-se o trabalho em Linguística, uma vez que análises quantitativas dependem de uma descrição qualitativa anterior e posterior à contagem. As abordagens qualitativas, por outro lado, se fiam, de maneira implícita, à ideia de quantidade, na medida em que consideram uma dada análise mais confiável se esta apresentar mais instâncias que a comprovem. Assim sendo, a autora afirma que “esses termos representam uma dicotomia de algum modo artificial” (SCHIFFRIN, 1987, p.66).

Embora a análise quantitativa esteja mais voltada às generalizações, e a qualitativa às particularidades, Schiffrin ressalta a importância de se combinar as duas abordagens uma vez que elas realizam contribuições complementares.

De tal modo, no trabalho que realizaremos nesta dissertação, se faz necessário alinhar as duas perspectivas, de maneira a considerar tanto a reiteração de determinados elementos e padrões, o que pode apontar para a proeminência de algum uso específico, quanto as peculiaridades de um caso particular, que podem contribuir na descrição e compreensão de um dado sob análise.

Assim como na FrameNet americana, em nossas análises usaremos como dados construções extraídas de *corpora* – tema de nossa próxima seção. À luz dessa discussão, vale ressaltar que as sentenças utilizadas constituem antes um elemento de representação, que um elemento para fins estatísticos. Em outros termos, as sentenças revelam o uso efetivo dos falantes, uma vez que são resultados de produções reais, e propõem-se a atestar esses usos, e não compor um cenário de probabilidades quanto aos padrões mais recorrentes.

1.2 Corpus

Nesta seção apresentaremos, primeiramente, o *corpus* utilizado nesta pesquisa. Em seguida, descreveremos os *softwares* empregados no tratamento de nosso *corpus*.

1.2.1 Constituição dos Corpora da FrameNet Brasil

A dissertação aqui apresentada faz parte do Projeto FrameNet Brasil (FN-Br), o qual será apresentado no próximo capítulo. Atualmente, a FN-Br realiza suas pesquisas através de dois *corpora* do português do Brasil (PB), o NURC-RJ e o Copa, acessíveis através do Sketch Engine – *software* que apresentaremos adiante.

O NURC-RJ é um *corpus* constituído por entrevistas gravadas entre as décadas de 1970 e 1990 – totalizando 350 horas – com informantes de nível superior completo, nascidos no Rio de Janeiro e filhos de pais,

preferencialmente, cariocas; possui 1.659.035 tokens. As transcrições do projeto NURC-RJ foram gentilmente cedidas à FN-Br por seus organizadores.

O *corpus* Copa foi criado para atender ao subprojeto da FN-Br “Copa 2014”. Composto de textos extraídos de *sites* governamentais de fomento ao turismo, *blogs* de viajantes e do *site* da FIFA, o Copa contém 623.088 tokens e continua a se expandir, à medida que novos textos são coletados e processados computacionalmente.

A FN-Br tem se dedicado à coleta de outros *corpora* a fim de ampliar sua base de dados, buscando abranger gêneros textuais diversificados. Os *corpora* do projeto contemplam dados de língua oral bem como de língua escrita. Em sua totalidade, os *corpora* acima descritos contêm 2.282.123 tokens.

Na implantação do Constructicon, optou-se por utilizar o NURC-RJ já que, por reunir dados de língua falada, tende a conter mais instâncias das construções da família *Para Infinitivo* – rede de construções escolhida para essa empreitada. Para aqueles casos em que uma dada construção dentre as definidas previamente por Torrent (2009; 2011; no prelo) não tivesse instanciações atestadas no NURC-RJ, recorreremos ao *corpus* Cetenfolha, também acessível via Sketch Engine, o qual é composto por cerca de 25 milhões de tokens.

1.2.2 *Tratamento Computacional*

Descreveremos, a seguir, o procedimento de tratamento computacional dado aos *corpora* desde o momento de sua coleta, até a anotação realizada segundo a metodologia da FrameNet.

Ao selecionar um determinado *corpus*, esse precisa ser depositado em um *software* que auxilie em consultas a esses dados, o Sketch Engine. Antes, contudo, o *corpus* selecionado precisa passar por um *parser*, ou seja, um analisador morfossintático.

O *parser* empregado nos *corpora* selecionados pela FN-Br, o PALAVRAS, foi originalmente projetado para a modalidade escrita do português brasileiro mas, atualmente, reconhece ortografia e gramática do português europeu (BICK, 2000). Foi utilizada a versão padrão de distribuição do *parser*, lembrando que esse recurso não foi projetado especificamente para

o trabalho aqui desenvolvido. Para os *corpora* da FN-Br, utiliza-se o módulo PAL-SYN do PALAVRAS, o qual, após identificar o lema de cada palavra, etiqueta-a quanto a sua classe morfológica, e quanto a sua função sintática. Uma vez que o *corpus* foi processado pelo *parser*, ele poderá ser utilizado em ferramentas de busca, tais como o Sketch Engine, que considerem os contextos sintáticos nos quais os lemas buscados se manifestam.

Depois de realizada a etiquetagem pelo PALAVRAS, o *corpus* precisa ser adaptado ao formato utilizado pelo Sketch Engine. Esse trabalho é feito pelo SYN-to-TAG, *software* desenvolvido pela FN-Br para desempenhar tal tarefa, o qual remove do *output* gerado pelo PALAVRAS, etiquetas que não são processáveis pela gramática do português brasileiro disponibilizada pelo Sketch Engine.

O Sketch Engine é utilizado pela FN-Br para busca e extração de *corpus*. Além de repositório de *corpus*, esse *software* constitui-se em “um sistema de consulta a *corpus* que incorpora esboços de texto, uma página, ou resumo” (cf. descrição presente no endereço eletrônico) e foi projetado para estudos a respeito do comportamento das palavras. O *corpus* já adaptado pelo SYN-to-TAG é, então, compilado no Sketch Engine.

No trabalho desenvolvido pela FN-Br, dois mecanismos de busca são utilizados no Sketch Engine: o *Word Sketch* e o *Concordance*. O primeiro deles fornece listas de palavras do *corpus* escolhido. Veja a Figura 1:



Figura 1: busca pelo *Word Sketch*

Na Figura 1, vemos uma imagem da página de busca do *Word Sketch*, na qual se consulta o lema do verbo “dar”. Na Figura 2, temos o resultado gerado a partir dessa busca:

Sketch Engine

About Home Settings Change password Log out

Search in Help

user: Prof. Maria Margarida Salomao corpus: NURC_RJ TAGGED Search in NURC_RJ TAGGED

Concordance
Word List
Word Sketch
Thesaurus
Find X
Sketch-Diff
?

Save
Change options
Clustering
Sorting
Gramrels
MW links

dar NURC_RJ TAGGED freq = 3524 (2124.1 per million)

modifier	1112	1.3	objeto	1930	4.6	sujeito	435	1.6
não	454	9.33	aula	132	10.91	gente	24	7.62
assim	59	8.64	trabalho	61	9.6	pai	6	7.56
só	32	8.53	idéia	37	9.0	professor	5	7.53
mais	48	8.43	impressão	32	8.87	dinheiro	6	7.51
então	30	8.37	condição	28	8.62	aspecto	4	7.46
bem	25	8.36	nome	33	8.62	carro	6	7.46
ai	19	8.12	um	39	8.58	governo	4	7.43
sempre	23	8.1	exemplo	21	8.44	nome	6	7.29

Figura 2: resultado da busca pelo *Word Sketch*

O resultado da busca pelo *Word Sketch* gera a página exibida pela Figura 2. As listas revelam os ambientes sintáticos em que o verbo “dar” figura no *corpus* buscado, o NURC-RJ. Desse modo, os diversos sujeitos que o verbo possui são listados (terceira coluna da Figura 2) junto a um número que indica a quantidade de sentenças em que cada sujeito diferente aparece. Esse mecanismo de busca lista, além do sujeito, objeto, modificadores (advérbios), e preposições pospostas ao verbo buscado.

O *Word Sketch* apresenta um resultado baseado em frequência. Conseqüentemente, se, por exemplo, uma determinada preposição aparecer em baixa frequência no *corpus* buscado, ela não constará no resultado da busca. Assim, se em uma consulta for preciso encontrar um ambiente em que “dar” apareça antecedido à preposição “para” – situação recorrente na família *Para Infinitivo* –, e o resultado não retornar nenhuma situação em que isso ocorra, será preciso realizar a busca por outro mecanismo, o *Concordance*.



Figura 3: busca pelo *Concordance*

Note-se que, na Figura 3, página de busca do *Concordance*, a consulta é feita com “dar para”. Como esse mecanismo não oferece listas distintas contendo os diferentes ambientes em que “dar” figura, mas sim, uma lista única abarcando todas as situações em que se encontra no *corpus* o elemento buscado, é preciso inserir o verbo “dar” seguido da preposição “para”, caso seja essa a disposição que se deseja obter no resultado. Veja a Figura 4:

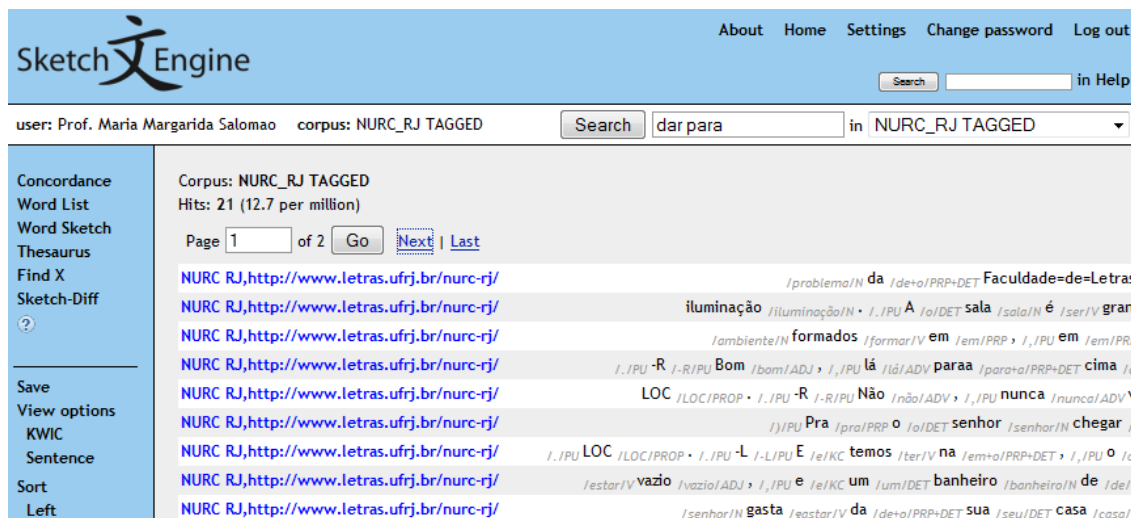


Figura 4: resultado da busca pelo *Concordance*

A Figura 4 exhibe o resultado encontrado da busca por “dar para”. Cada linha iniciada pelo endereço da fonte do *corpus* contém uma sentença.

De posse das sentenças, essas serão importadas para o Desktop através do *site* FN-Br Admin. No Desktop as sentenças extraídas dos *corpora* serão anotadas conforme a metodologia da FrameNet (RUPPENHOFER et al.,

2010). As sentenças anotadas servem de exemplo para os diversos *frames*, Unidades Lexicais e construções¹ descritos no Desktop.

Segundo a anotação proposta pela FrameNet, as sentenças são classificadas quanto à sua valência sintática e semântica. A valência sintática diz respeito à Função Gramatical e Tipo Sintagmático dos itens presentes na sentença, enquanto a valência semântica considera os Elementos de *Frame* evocados pela Unidade Lexical alvo. Todo o processo de anotação lexicográfica será detalhado na subseção 1.3.1.

O Desktop realiza, ainda, o agrupamento das sentenças em padrões de valência. Esses padrões são instituídos em função do comportamento gramatical e sintagmático que cada sentença apresenta em relação a seus Elementos de *Frame*.

1.3 Métodos de Anotação na FrameNet Brasil

Depois de receber o tratamento computacional anteriormente apresentado, as sentenças colhidas nos *corpora* são submetidas à anotação no Desktop. Atualmente, a FrameNet possui três diferentes métodos de anotação, quais sejam, o lexicográfico, o de texto corrido e o de construções.

1.3.1 Anotação Lexicográfica

O primeiro método de anotação desenvolvido pela FrameNet, o lexicográfico, tem como centro a Unidade Lexical (UL). Ao se anotar lexicograficamente, o que se faz, na verdade, é realizar uma anotação a partir de uma UL alvo. Todas as análises, portanto, serão feitas em termos de quais Elementos de *Frame* tal UL perfila, ou qual *frames* ela evoca.

Como a anotação lexicográfica tem por meta anotar as diversas situações em que a UL figura, se um dado lema a ser anotado evocar mais de um *frame*, será anotado separadamente em cada *frame* evocado, exibindo-se as distribuições sintático-semânticas encontradas em cada um deles. É preciso

¹ Os termos “*frames*”, “Unidades Lexicais” e “construções” serão definidos no capítulo 2. Por agora, cabe apontar que uma Unidade Lexical é o pareamento de um lema a um *frame* específico.

fazê-lo, uma vez que a anotação se modifica conforme o *frame* evocado. Diferentes papéis serão atribuídos aos constituintes da sentença, dependendo do *frame* que a UL evoca.

Esse modo de anotação respeita a localidade sintática, o que significa que só será anotada a sentença na qual a UL se encontra. Qualquer outro termo que não esteja no escopo da sentença, ou que tenha sido omitido e seja recuperável pelo contexto é considerado conforme o que a sentença da UL alvo expõe.

A anotação lexicográfica é realizada, conforme a descrição exposta no Book (RUPPENHOFER et al., 2010), através de camadas, as quais se decompõem em: Elemento de *Frame*, Função Gramatical, Tipo Sintagmático, uma camada extra dependente da classe de palavras na qual se enquadra a unidade sendo anotada, a qual será esclarecida em breve, e a camada Outros.

A camada Elemento de *Frame* (EF) é usada para etiquetar os constituintes da sentença conforme os papéis semânticos que assumem em relação à UL, assinalando os nomes dos EFs nas cores correspondentes às que foram marcadas quando da criação do EFs – essas cores serão mantidas no decorrer da análise para facilitar sua identificação.

As classes gramaticais de cada constituinte da sentença alvo serão assinaladas na camada Função Gramatical (FG). A FrameNet classifica os sintagmas e palavras codificadores dos EFs em uma das funções a seguir: externo (Ext), o EF que atua como sujeito na sentença; objeto direto (DObj); objeto indireto (IndObj); dependente (Dep), função que abarca casos de adjunção; além de aposto (Aposto) e núcleo (Núcleo), funções utilizadas quando as palavras alvo são nomes ou adjetivos.

A terceira camada contempla o Tipo Sintagmático (TS) no qual se encaixa cada EF. A classificação ocorre em função do sintagma em que um dado constituinte se encontra, por exemplo, se verbal, ele é marcado como SV, se nominal, como SN.

É possível contarmos, ainda, com uma camada extra coindexada à classe de palavras à qual a UL alvo pertence. Assim, se essa tratar-se de um verbo, a camada receberá o nome de Verbo. A essa camada adiciona-se, quando necessário, informações relacionadas à estrutura da UL.

A anotação é concluída na camada Outros, onde são compreendidas quaisquer informações referentes à UL não captadas anteriormente pelas outras camadas. Tal camada é usada, especialmente, para a marcação de estruturas com verbos suporte e cópula, além de orações relativas.

1.3.2 Anotação de Texto Corrido

Diferentemente da versão lexicográfica, a anotação de texto corrido não possui uma UL alvo a partir da qual a anotação seja exclusivamente realizada. Ao contrário, seleciona-se um texto e todos os constituintes nele encontrados capazes de evocar um *frame* são anotados. Em outras palavras, na anotação de texto corrido, cada UL encontrada é anotada.

Dessa forma, segue-se a sequência que o texto apresenta. Não há um *frame* único a direcionar a anotação; como é a UL que evoca o *frame* e são diversas as ULs a se anotar, diversos serão, também, os *frames* a se anotar. Conforme aponta Fillmore (2008a), nessa versão, a anotação de cada UL da sentença é relevante para que se possa organizar a estrutura semântica da sentença como um todo.

A anotação de texto corrido constitui-se em vários “conjuntos de camadas, cada uma correspondendo a uma UL alvo” (FILLMORE, 2008a). Torrent (2012) assinala que as ULs são anotadas de modo a criar “árvores de dependência que relacionam as ULs – e seus papéis em cada um dos frames evocados na sentença – entre si”. Assim sendo, não é possível ter uma visão completa da anotação de uma sentença longa, mas já existem trabalhos sendo feitos a fim de otimizar sua visualização.

Observe-se no Quadro 1 um exemplo de anotação de texto corrido. A anotação decorre de uma monografia em que foram selecionadas redações para se realizar tal anotação com o objetivo de se verificar a adequação de redações de alunos do Ensino Médio aos critérios propostos pelo ENEM (SILVA & ESTRADA, 2012):

Horário Eleitoral

1a. [O horário eleitoral obrigatório_{Entidade}] é_{cop} **bom**_{Usefulness} [para o povo e bom também para a democracia_{Finalidade}].

2a. [Nele_{Meio}], [a população_{Agente}] tem_{Sup} a [**oportunidade**_{Oportunidade}]**Opportunity** [de conhecer melhor os candidatos e suas propostas_{Situação_desejável}].

2b. Nele, [a população_{Conhecedor}] tem a oportunidade de **conhecer**_{Familiarity} [melhor_{Grau}] [os candidatos e suas propostas_{Entidade}].

3a. [Pesquisa do Datafolha_{Entidade}] **mostra**_{Cause_to_perceive} [que 64% da população do Estado de São Paulo acha necessária a manutenção do horário eleitoral_{Fenômeno}].

3b. Pesquisa do Datafolha mostra que [64% da população do Estado de São Paulo_{Pensador}] **acha**_{Opinion} [necessária a manutenção do horário eleitoral_{Parecer}].

4a. Ao se **conceder**_{Giving} [espaço nas emissoras de tv's_{Tema}] [para os partidos políticos_{Recipiente}], o povo tem a oportunidade de ver e analisar melhor as propostas de seus candidatos. [INC=Doador]

4b. [Ao se conceder espaço nas emissoras de tv's para os partidos políticos_{Condição}], [o povo_{Agente}] tem_{sup} a [**oportunidade**_{Oportunidade}]**Opportunity** [de ver e analisar melhor as propostas de seus candidatos_{Situação_desejável}].

4c. [Ao se conceder espaço nas emissoras de tv's para os partidos políticos_{Evidência}], [o povo_{Pensador}] tem a oportunidade de **ver**_{Grasp} e analisar [melhor_{Maneira}] [as propostas de seus candidatos_{Fenômeno}].

4d. [Ao se conceder espaço nas emissoras de tv's para os partidos políticos_{Evidência}], [o povo_{Avaliador}] tem a oportunidade de ver e **analisar**_{Assessing} [melhor_{Maneira}] [as propostas de seus candidatos_{Fenômeno}].

5a. No entanto, [algumas mudanças_{Ação}] devem ser_{cop} **feitas**_{Intentionally_act} [a fim de se obter um horário eleitoral mais justo_{Finalidade}]. [INC=Agente]

5b. No entanto, algumas mudanças devem ser feitas a fim de se **obter**_{Getting} [um horário eleitoral mais justo_{Tema}]. [INC=Recipiente]

6a. [O tempo de visibilidade_{Todo}], por exemplo, deve ser_{cop} **distribuído**_{Separating} [de forma igualitária_{Modo}] [para todos os candidatos e coligações_{Recipientes}]. [IND=Partes], [INC=Agente]

7a. Deve-se, ainda, **exigir**_{Being_obligated} [que os candidatos_{Parte_responsável}] [usem seu tempo para os debates e apresentação de propostas_{Dever}], [evitando-se assim o uso de propagandas_{Consequência}].

<p>7c. Deve-se, ainda, exigir que [os candidatos_{Agente}] usem_{Using} [seu tempo_{Instrumento}] [para os debates e apresentação de propostas_{Finalidade}], [evitando-se assim o uso de propagandas_{Resultado}].</p> <p>7d. Deve-se, ainda, exigir que os candidatos usem seu tempo para os debates e apresentação de propostas, evitando-se_{Avoiding} [assim_{Maneira}] [o uso de propagandas_{Situação_indesejável}]. [INC=Agente]</p> <p>8a. Outro ponto importante é_{cop} a proibição_{Prohibiting} [da divulgação de pesquisa eleitoral por parte dos candidatos_{Estado_de_coisas}]. [IND=Princípio]</p> <p>9a. [O eleitor_{Protagonista}] deve guiar-se_{Compliance} [somente_{Maneira}] [pela análise de seu candidato_{Norma}], não sendo influenciado por pesquisas, que em alguns casos são mal intencionadas.</p> <p>9b. [O eleitor_{Entidade_dependente}] deve guiar-se somente pela análise de seu candidato, não sendo_{Cop} influenciado_{Objective_Influence} [por pesquisa_{Entidade_influenciadora}], [que em alguns casos são mal intencionadas_{Descritor}].</p> <p>10a. Portanto, [o horário eleitoral obrigatório_{Ator}] constitui_{Performers_and_roles} [um importante mecanismo para o sistema democrático_{Papel}].</p> <p>11a. [Ele_{Ajudador}] ajuda_{Assistence} [na divulgação de informação_{Meta}] e auxilia o cidadão na escolha de seu.</p> <p>11b. [Ele_{Auxiliador}] ajuda na divulgação de informação e auxilia_{Assistence} [o cidadão_{Beneficiário}] [na escolha de seu candidato_{Meta}].</p> <p>12a. Dessa forma, tem-se_{Coming_to_believe} [que o horário eleitoral é bom para o eleitor_{Conteúdo}]. [INC=Pensador], [IND=Evidência]</p> <p>12b. Dessa forma, tem-se que [o horário eleitoral_{Entidade}] é_{cop} bom_{Usefulness} [para o eleitor_{Finalidade}].</p>

Quadro 1: Anotação de texto corrido

A numeração indica cada sentença analisada. As letras sinalizam as sentenças anotadas mais de uma vez por possuírem mais de um constituinte evocador de *frame*. A segunda sentença do texto, por exemplo, possui duas ULs – em negrito –, a saber, “oportunidade” e “conhecer”; por isso tem-se 2a e 2b. Para cada UL, uma anotação é realizada separadamente.

Ao lado de cada UL vem anotado o nome do *frame* por ela evocado, como em 2b, em que a UL “conhecer” aparece seguida de *Familiarity*, *frame* de Familiaridade. Os Elementos de *Frame* são amparados por colchetes e a eles

se segue o nome que recebem, como na sentença 2a, em que “a população” é o Elemento de *Frame* “Conhecedor”, “melhor” é Grau, e “os candidatos e suas propostas” é Entidade. Os Elementos de *Frame* são estabelecidos em função do *frame* evocado pela UL.

1.3.3 Anotação Construcional

A anotação construcional aproxima-se da anotação lexicográfica na medida em que escolhe um item alvo a ser buscado e anotado. Se na versão lexicográfica podemos escolher como UL alvo o verbo “dar”, no modo construcional é possível escolher a Construção Modal Epistêmica como alvo a ser anotado.

Por outro lado, a anotação construcional aproxima-se da anotação de texto corrido uma vez que não se prende apenas à unidade alvo, como ocorre na anotação lexicográfica. O que for relevante para a descrição de tal unidade alvo, mesmo que se trate de um dado do contexto, será incorporado à anotação.

Na anotação construcional, o que anotamos, na verdade, são as propriedades de um construto. De modo paralelo ao que vimos na anotação lexicográfica, assim como os Elementos de *Frame* compõem a cena evocada pela UL, os elementos que compõem a construção são denominados Elementos da Construção.

A camada do Tipo Sintagmático apresentada na descrição da anotação lexicográfica se mantém. Entretanto, a camada da Função Gramatical não é obrigatória na anotação construcional.

Duas novas camadas são acrescentadas à anotação construcional, a do Tipo Sintagmático do Construto e a do Elemento Evocador da Construção (EEC). A primeira indica o tipo de sintagma do construto alvo, e não apenas de suas partes componentes – como ocorre com a camada do Tipo Sintagmático. A camada do EEC, por sua vez, aponta a configuração responsável por evocar o padrão construcional alvo.

2 FRAMES E CONSTRUÇÕES

Neste capítulo apresentaremos os aportes teóricos que norteiam a presente dissertação. Assim, iniciaremos pela Semântica de *Frames*, passando pela Gramática das Construções e, finalmente, o Constructicon.

2.1 Semântica de *Frames*

Nesta seção abordaremos inicialmente a Semântica de *Frames*, seu surgimento, propriedades e principais categorias analíticas. Em seguida, apresentaremos a FrameNet, a instanciação mais sofisticada da Semântica de *Frames* do ponto de vista das abordagens da Linguística Computacional e da Linguística de *Corpus*.

2.2.1 Origem e Principais Desdobramentos

A Linguística moderna, embora seja um campo investigativo recente – se comparado, por exemplo, às ciências naturais –, possui uma história permeada por grandes mudanças de paradigmas. No início do século XX, Ferdinand de Saussure, em sua obra póstuma *Cours de Linguistique Générale* (2008 [1916]), confere um caráter formal à disciplina. O livro que apresenta as quatro tão discutidas dicotomias saussureanas – *langue X parole*, sincronia X diacronia, sintagma X paradigma, significado X significante – torna-se um marco seminal de estudos estruturalistas na Linguística.

Na segunda metade do século XX, os estudos de Noam Chomsky vêm fundar a Linguística Gerativa, segundo a qual, a língua é tratada sob seu aspecto formal, de modo a ser manipulada por um sistema de regras capazes de **gerar** infinitas sentenças. Chomsky (1965) traça uma distinção entre competência e desempenho, em que a primeira refere-se ao conhecimento linguístico interno que permite ao falante produzir sentenças em sua língua, e o segundo alude ao comportamento linguístico, o qual resulta da competência bem como de fatores extralinguísticos. Ao caracterizar a competência, Chomsky assinala o caráter inato da linguagem, ou seja, uma capacidade adquirida por meio genético, própria da espécie humana. Nesse ponto, é

possível localizar essa corrente teórica como cognitivista em razão de a língua ser entendida como um fenômeno mental.

Como a língua era vista sob seu aspecto formal, a sintaxe era o campo privilegiado no Gerativismo. Assim, em decorrência da insatisfação de alguns estudiosos com a Semântica Gerativa (cf. SALOMÃO, 2009b), a qual se mostrava insuficiente para abarcar a complexidade das relações entre forma e sentido, entre a sintaxe e a semântica, surgem os primeiros trabalhos em Linguística Cognitiva. Conforme aponta Salomão (2007a), essa vertente emergirá através de algumas proposições, tais como a de George Lakoff, segundo a qual categorias linguísticas sofrem efeitos de prototipia assim como categorias cognitivas e culturais, e a de Charles Fillmore acerca da Semântica de *Frames*, a qual resulta de “seus esforços anteriores para postulação de uma ‘gramática de casos’” (SALOMÃO, 2007a, p.33). Passemos, nesse momento, a uma breve retrospectiva de alguns aspectos presentes nesse trabalho, a fim de avaliarmos em que medida os primórdios da Semântica de *Frames* estão nele contidos.

Em *The Case for Case*, Fillmore (1968), proveniente do grupo de semanticistas gerativos, promove um estudo do sentido, em que trabalha os papéis que cada sintagma assume em uma dada sentença em relação a seu predicador. Para tanto, ele usa um número limitado de relações de caso, como, por exemplo, Agente, Paciente e Instrumento, os quais identificamos nos exemplos de (1) a (3), semelhantes aos que o autor sugere no referido texto:

- (1) [Alice _{Agente}] quebrou [a janela _{Paciente}]
- (2) [O martelo _{Instrumento}] quebrou [a janela _{Paciente}]
- (3) [Alice _{Agente}] quebrou [a janela _{Paciente}] [com um martelo _{instrumento}]

No primeiro exemplo, o sujeito cumpre papel de Agente em relação ao verbo, ao passo que, no segundo exemplo, seu papel passa a ser o de Instrumento. No terceiro exemplo, os dois papéis aparecem na sentença, mas a posição de sujeito é preenchida com o papel de Agente.

Observe o exemplo (4):

- (4) *Alice e o martelo quebraram a janela.

O fato de o exemplo (4) não ser aceitável se explica devido à diferença existente entre os dois sujeitos no que tange ao seu papel semântico, conforme observamos nas duas primeiras sentenças. Em outras palavras, para Fillmore, a agramaticalidade dessa sentença se justificaria pelo fato de que apenas sintagmas nominais com o mesmo caso podem ser coordenados.

Em *The Case for Case* se observa a origem de conceitos primitivos que, posteriormente, serão empregados na Semântica de *Frames*, apesar de a própria noção de *frame*, a ser discutida mais adiante, ainda não havia sido trazida à Linguística. Dessa forma, é possível afirmar que esse trabalho de Fillmore é precursor da Semântica de *Frames*, na medida em que reconhece a propriedade que os itens lexicais têm de organizar os argumentos de uma sentença. Pode-se afirmar, ainda, que nele se encontram as bases para a Gramática das Construções, uma vez que, conforme se viu nos exemplos, assume-se que um mesmo item lexical pode aparecer em estruturas argumentais diferentes, as quais se organizam segundo as relações de caso atribuídas a cada argumento que as compõem.

A Linguística Cognitiva surge, portanto, no esteio de estudos que, como o trabalho de Fillmore (1968), nascem de um descontentamento com o tratamento dado à semântica pelo Gerativismo. Desse modo, na década de 1980, a Linguística Cognitiva começa a se delinear como novo paradigma teórico. Ainda que ambas possam ser chamadas de áreas de investigação cognitivistas, essa nova vertente diferencia-se do Gerativismo, na medida em que reconhece a língua como resultado da interação de diversas capacidades cognitivas, assim como de fatores biológicos e socioculturais. No Gerativismo, a língua é vista como uma faculdade modular no cérebro, independente de outras faculdades. Geeraerts e Cuyckens esclarecem a diferença no escopo investigativo dos dois campos teóricos:

A Linguística Cognitiva está interessada em nosso conhecimento de mundo e estuda a questão de como a linguagem natural contribui para ele. O linguista gerativo, ao contrário, está interessado em nosso conhecimento da língua e pergunta como tal conhecimento pode ser adquirido dada uma teoria cognitiva de aprendizagem (GEERAERTS & CUYCKENS, 2007, p.6)².

² Todas as traduções presentes nesta dissertação são de responsabilidade da autora. "Cognitive Linguistics is interested in our knowledge of the world and studies the question how natural language contributes to it. The generative linguist, conversely, is interested in our knowledge of the language and asks the question how such knowledge can be acquired given a cognitive theory of learning."

Para Salomão, as questões a serem salientadas como definidoras da Linguística Cognitiva são as seguintes:

a ênfase na natureza cognitiva das categorias linguísticas e dos processos gramaticais (...); o reconhecimento da diversidade das línguas do mundo e, mesmo, dos usos de uma mesma língua, como condição constitutiva da linguagem humana, que é também produção social e histórica (...); o tratamento processual da significação, incorporando seu inerente pragmatismo através dos conceitos centrais de perspectiva, enquadramento e espaços mentais (...); a reivindicação do estudo dos processos imaginativos (da metáfora, da metonímia, da contrafactualidade, da mesclagem) como característica distintiva de cognição e de linguagem humana (SALOMAO, 2009b).

Para lidar com o significado, a teoria gerativa se vale das condições de verdade sob as quais uma sentença pode ser julgada. Assim, conforme aponta Ferrari (2011, p.14-15), o significado é concebido como reflexo da realidade, ao passo que, para a Linguística Cognitiva, trata-se de uma “construção cognitiva através da qual o mundo é apreendido e experienciado”. O significado é, portanto, o campo privilegiado da Linguística Cognitiva, e é nesse contexto, em contraposição à semântica vericondicional, que surge uma nova proposta para o estudo do sentido.

Fillmore (1985) propõe uma discussão a respeito da Semântica da Verdade (*T-semantics*³) – referindo-se, aqui, às teorias semânticas baseadas em julgamentos de verdade – em oposição à Semântica do Entendimento (*U-semantics*⁴) – que reúne teorias semânticas fundamentadas na compreensão da língua. Para tanto, o linguista considera uma sentença distinguindo os tipos de análise que seriam feitos a partir dessas duas abordagens, qual seja: “*My dad wasted most of the morning on the bus*⁵”.

Em uma análise sob a ótica da Semântica da Verdade, estaremos interessados nas condições de verdade a que essa elocução se submete. Assim, a sentença só é verdadeira se a pessoa mencionada no sintagma nominal “*my dad*” (meu pai) tiver, de fato, gastado a maior parte da manhã no ônibus referido pelo sintagma “*the bus*” (no ônibus), na parte do dia aludida pelo sintagma “*the morning*” (da manhã). Se meu pai gastou a maior parte da

³ Semantics of Truth.

⁴ Semantics of Understanding.

⁵ Meu pai desperdiçou a maior parte da manhã no ônibus.

manhã no ônibus, então essa sentença (falada por mim, hoje) é verdadeira (FILLMORE, 1985, p. 230-231).

A Semântica do Entendimento, por sua vez, estará interessada no uso da palavra “*dad*” (termo informal para “pai”) ao invés de “*father*” (pai), que revela o tipo de relação entre o falante e seu pai; no fato de que “*my dad*” foi usado em vez de simplesmente “*dad*”, o que aponta para a relação entre o falante e o endereçado – os quais não são membros da mesma família; no uso da palavra “*wasted*” (“desperdiçou”) em vez de “*spent*” (grosso modo, “passou”), que mostra o julgamento por parte do falante de que o tempo não foi empregado de maneira proveitosa (FILLMORE, 1985, p. 231).

Desse modo, conforme Fillmore assinala, a Semântica do Entendimento assume que categorias codificadas linguisticamente pressupõem crenças sobre o mundo, experiências partilhadas, padrões de como fazer as coisas e maneiras de ver as coisas. Assim sendo, para o autor, os significados podem ser explicados a partir de *frames*, estruturas complexas de conhecimento fundamentadas em expectativas partilhadas socialmente. Os *frames* se constituem num sistema de conceitos relacionados de tal modo que, “para entender um deles, é necessário entender toda a estrutura na qual ele se encaixa” (FILLMORE, 1982).

Os *frames* fazem parte da vida do ser humano nas cenas mais corriqueiras do cotidiano. Saber que o comportamento que se tem num *happy hour* em um bar é diferente do que é encontrado numa palestra envolve compreender a que *frame* cada uma dessas situações pertence. Entender que o esquema evocado por “falar” é diferente do que é evocado por “gritar” requer conhecimento adquirido através da experiência.

Nosso conhecimento sobre o mundo é organizado e se expressa linguisticamente. A fim de clarificar tais apontamentos, observe a charge na Figura 5.



Figura 5: Charge ⁶

Para compreendermos essa charge, é necessário entender quais os *frames* evocados por “mãe” e “madrasta”. Enquanto “mãe”, nesse caso, evoca a ideia de um ser que age com piedade e paciência, que facilita os percalços da vida, “madrasta” evoca justamente a ideia oposta, a de um ser capaz de criar todos os obstáculos possíveis a fim de atormentar a vida do indivíduo. Assim, a Polícia Federal seria essa entidade que causa problemas ao detento, ao passo que a Justiça, isto é, o judiciário, representaria a entidade facilitadora.

Dessa forma, a Semântica de *Frames* se instaura como um programa de pesquisa em semântica empírica que enfatiza o *continuum* entre linguagem e experiência (PETRUCK, 1996). A cena da Transação Comercial, descrita por Fillmore (1982), elucida de que modo um *frame* estrutura os significados emergentes numa dada cena. Nesta cena há uma pessoa interessada na troca de dinheiro por mercadorias (Comprador), outra interessada na troca de mercadorias por dinheiro (Vendedor), as mercadorias que o Comprador pode adquirir (Mercadorias), e o dinheiro a ser adquirido pelo Vendedor (Dinheiro). A partir dessa cena, é possível observar que há verbos que se relacionam semanticamente, cada um evocando uma mesma cena geral. Fillmore propõe que:

⁶ Disponível em: http://www.glimboo.com/imagens_charges.php?p=6

o verbo *comprar* foca nas ações do Comprador em relação às Mercadorias, deixando em segundo plano o Vendedor e o Dinheiro; que o verbo *vender* foca nas ações do Vendedor em relação às Mercadorias, colocando em segundo plano o Comprador e o Dinheiro; que o verbo *pagar* foca nas ações do Comprador em relação ao Dinheiro bem como ao Vendedor, deixando em segundo plano as Mercadorias, e assim por diante, com verbos tais como *gastar*, *custar*, *cobrar* e um número de outros que são, de algum modo, mais periféricos que esses (FILLMORE, 1982, p.378).⁷

Dois termos precisam ser esclarecidos, a saber, *evocar* e *invocar*. Ao ler um texto, os elementos linguísticos *evocam* os *frames*, de modo a instanciar uma cena na mente do leitor. Esse leitor, por sua vez, *invoca* um *frame* ao atribuir uma interpretação a um dado texto. Dessa maneira, a sentença “Logo que todos terminaram de cantar, Maria soprou as velhinhas e fez um pedido” não contém nenhuma referência explícita ao aniversário de Maria, no entanto, um leitor que compartilhe conhecimentos culturais é capaz de *invocar* a cena de aniversário.

Os *frames* podem variar em diferentes culturas. No ocidente, por exemplo, o luto está relacionado à cor preta e, por isso, é comum que as pessoas se vistam de tal cor nessa ocasião. No oriente, porém, a cor relativa ao luto é branca. Em uma temática relacionada, enquanto na cultura brasileira o *frame* do Dia de Finados envolve, em especial nas comunidades mais tradicionais do interior, um comportamento recatado e introspectivo; na cultura mexicana ele é marcado por festividades coloridas e alegres.

De tal modo, muito além de proposições estáticas, sem qualquer interferência da cultura, Fillmore explica que os *significados são relativizados a cenas* (1977). Salomão (2009a) pontua que esse princípio carrega em seu âmago a concepção de um *continuum* entre linguagem e experiência – relação essencial à composição da Linguística Cognitiva.

Nesse contexto, como elemento fundamental, a Semântica de *Frames* abre espaço para um projeto em lexicografia que busca alinhar a Linguística Cognitiva à descrição da língua na perspectiva computacional. É sobre esse assunto que trataremos na próxima seção.

⁷ “(...) the verb *buy* focuses on the actions of the Buyer with respect to the Goods, backgrounding the Seller and the Money; that the verb *sell* focuses on the actions of the Seller with respect to the Goods, backgrounding the Buyer and the Money; that the verb *pay* focuses on the actions of the Buyer with respect both to the Money and the Seller, backgrounding the Goods, and so on, with such verbs as *spend*, *cost*, *charge*, and a number of others somewhat more peripheral to these.”

2.1.2 FrameNet

A FrameNet é um projeto em lexicografia computacional que extrai informação sobre as propriedades semântico-sintáticas de palavras do inglês a partir de variados *corpora* eletrônicos. Parte do trabalho é realizada manualmente, como na anotação das sentenças, e outra parte é desempenhada automaticamente por *softwares* (FILLMORE, JOHNSON & PETRUCK, 2003).

O projeto é sediado no International Computer Science Institute (ICSI) e é liderado por Charles Fillmore, da Universidade da Califórnia, em Berkeley. O nome do projeto evidencia sua relação com a Semântica de *Frames*, empregada como base teórica, além de seu interesse em redes de significados compostas pelas palavras e pelos *frames* que estruturam seus significados.

Por ser amparada por evidência em *corpora*, a FrameNet se diferencia de dicionários comumente encontrados no Brasil, uma vez que as sentenças disponibilizadas refletem o uso real do falante, não sendo meros artigos ilustrativos. Além disso, tais sentenças embasam a descrição sintático-semântica que é realizada. O projeto diferencia-se, ainda, dos *thesauri* e da WordNet, os quais organizam as palavras segundo relações de hiponímia, hiperonímia, sinonímia, antonímia, entre outras, sem, entretanto, considerar o *frame* como estrutura organizadora das relações entre os itens lexicais.

O trabalho na FrameNet consiste em identificar e descrever *frames*, analisar o significado das palavras a partir dos *frames* que evocam, e estudar as propriedades sintáticas das palavras questionando de que modo propriedades semânticas se manifestam na forma sintática (FILLMORE, JOHNSON & PETRUCK, 2003).

A descrição dos *frames* bem como outros trabalhos realizados pela FrameNet são disponibilizados em uma plataforma online⁸. Em seu banco de dados, já foram descritas mais de 11.600 Unidades Lexicais, das quais quase 7.000 estão completamente anotadas, em mais de 960 *frames*, exemplificados em mais de 170.000 sentenças.

⁸ Disponível em: <https://framenet.icsi.berkeley.edu/fndrupal/>

Inicialmente, a FrameNet se desenvolveu como recurso lexical para tratar da língua inglesa. Contudo, hoje já existem trabalhos em diversas línguas como o espanhol, o japonês, o alemão e o sueco. Em 2008, o projeto foi trazido pela professora Margarida Salomão para pesquisa em língua portuguesa, recebendo o nome de FrameNet Brasil (SALOMAO, 2009a). Desde a sua fundação na UFJF, a FrameNet Brasil conta com a parceria da professora Rove Chishman (UNISINOS), coordenadora do Projeto FrameCorp, o qual aplica a Semântica de *Frames* na anotação de *corpora*, também utilizando o banco de dados da FrameNet americana. O projeto brasileiro, do qual a presente dissertação faz parte, desenvolveu sua plataforma online⁹ a fim de disponibilizar dados e compartilhar projetos e resultados.

Duas são as unidades básicas de análise: o *frame* e a **Unidade Lexical** (UL), a qual consiste no pareamento de um lema a um *frame*. Se pensarmos no lexema “frio”, podemos relacioná-lo ao *frame* de Escala_de_Temperatura – em, por exemplo, “essa comida está fria” –, mas também é possível relacioná-lo ao *frame* de Avaliação_da_Interação_Social – como na sentença “Ele é sempre muito frio em seus comentários”. Assim, “frio” é capaz de evocar esses dois *frames*; em cada caso, temos uma Unidade Lexical diferente para o lexema adjetivo “frio”.

Outro componente importante a ser definido e utilizado na descrição dos *frames* é o **Elemento de Frame** (EF). Ao definirmos um *frame*, é possível identificar palavras que se relacionam à UL evocadora desse *frame*, cada uma desempenhando um papel distinto. Tais palavras, às quais um papel semântico é atribuído, são chamadas de Elementos de *Frame*, uma vez que contribuem com informações a respeito do *frame* que é evocado.

É importante assinalar que, na FrameNet, os *frames* são compostos de funções microtemáticas, os Elementos de *Frame*. Salomão (2009a) explica que os Elementos de *Frame* correspondem aos Casos Semânticos propostos por Fillmore na Gramática de Casos (FILLMORE, 1968). A diferença entre as abordagens é descrita pelo próprio Fillmore da seguinte forma:

O que distingue a Semântica de *Frames* da Gramática de Casos é a troca de se tomar os conceitos de papel como primários e os tipos de situação como definidos em termos dos papéis, por tomar o *frame* como primário e

⁹ Disponível em: <http://www.framenetbr.ufjf.br/>

definir distinções de papel de modo relativo ao *frame*¹⁰ (FILLMORE, 2008a, p.6).

Os Elementos de *Frame* são Funções Microtemáticas “postuladas em relação ao *frame* a que se referem” (SALOMÃO, 2009a). Assim, retomando o *frame* que vimos na seção anterior, o de Transação Comercial, alguns de seus Elementos de *Frame* são definidos como Vendedor e Comprador, e não como Agente, como seria classificado em uma abordagem mais tradicional. Veja no Quadro 2 a definição dada pela FrameNet do referido *frame*.

Commercial_transaction	
Definition:	
These are words that describe basic commercial transactions involving a Buyer and a Seller who exchange Money and Goods . The individual words vary in the frame element realization patterns. For example, the typical patterns for the verbs buy and sell are: BUYER buys GOODS from the SELLER for MONEY. SELLER sells GOODS to the BUYER for MONEY.	
His \$20 TRANSACTION with Amazon.com for a new TV had been very smooth.	
FEs:	
Core:	
Buyer [Byr]	The Buyer wants the Goods and offers Money to a Seller in exchange for them.
Goods [Gds]	The FE Goods is anything (including labor or time, for example) which is exchanged for Money in a transaction.
Money [Mny]	Money is the thing given in exchange for Goods in a transaction.
Seller [Slr]	The Seller has possession of the Goods and exchanges them for Money from a Buyer .
Non-Core:	
Means [Mns]	The means by which a commercial transaction occurs.
Semantic Type: State_of_affairs	
Rate [Rate]	Price or payment per unit of Goods.

¹⁰ What distinguishes frame semantics from case grammar is the switch from taking the role concepts as primary and the situation types as defined in their terms, to taking the frame as primary and defining role distinctions relatively to the frame.

Unit [Unit] The Unit of measure of the Goods according to which the exchange value of the Goods (or services) is set. Generally, it occurs in a by-PP.

Frame-frame Relations:

Inherits from: [Reciprocity](#)

Is Inherited by:

Perspective on:

Is Perspectivized in:

Uses:

Is Used by:

Subframe of: [Commerce scenario](#)

Has Subframe(s): [Commerce goods-transfer](#), [Commerce money-transfer](#)

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

Lexical Units: *transaction.n*

Quadro 2: Definição do *frame* de Transação_Comercial pela FrameNet

Note-se que, primeiramente, a definição do *frame* em questão é proposta. Em seguida, uma lista de seus Elementos de *Frames* é apresentada – à definição e aos EFs podem-se seguir sentenças exemplos. Cada EF (em inglês, *FE*) é marcado com fundo em cor distinta e contém uma definição própria. São exibidas as relações entre *frames* – sobre as quais discutiremos a diante – e, por fim, as Unidades Lexicais que evocam tal *frame* são listadas.

O problema em classificações genéricas – como Agente, em vez de Vendedor – é que pode-se etiquetar do mesmo modo elementos muito diferentes, o que gera imprecisão na descrição. Salomão esclarece que:

a proeminência epistemológica da noção de *frame* como categoria cognitiva torna desejável que os Elementos dos *Frames* sejam identificados a partir das específicas situações conceptuais de que participam. Caberá às **relações verificadas entre os *frames*** (de Herança, de Pressuposição, de Perspectiva...) representar as generalidades prevalentes entre os elementos dos esquemas conceptuais (SALOMÃO, 2009a, p.4) [grifo original].

Conforme Salomão assinala, a recuperação dos papéis abstratos relativos aos *frames* pode ser feita através do FrameGrapher, ferramenta que

disponibiliza a visualização das relações entre os *frames*, o que pode ser visualizado na Figura 6:

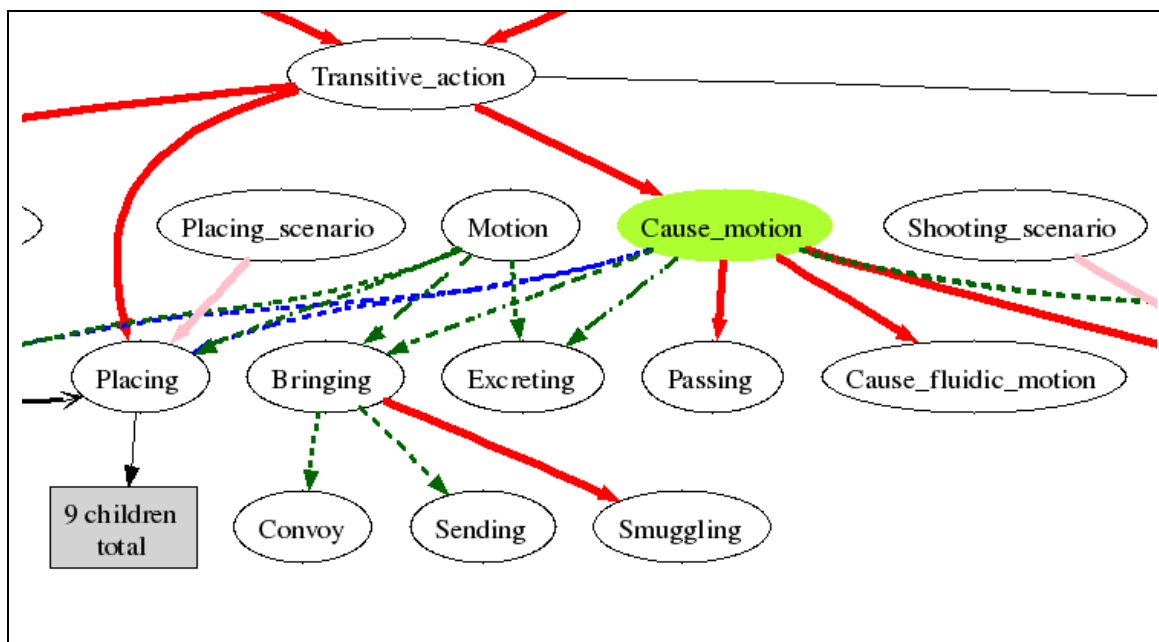


Figura 6: FrameGrapher: Cause_motion

Pesquisando a descrição do *frame* de *Sending* (Enviar), presente na Figura 6, encontraremos o Elemento de *Frame Sender* (Remetente) descrito como aquele que inicia o movimento do *Theme* (Tema – outro Elemento de *Frame*). Observando a Figura 6, nota-se que tal *frame* está conectado ao de *Bringing* (Portar) por uma relação de Uso. Esse *frame*, por sua vez, possui o Elemento de *Frame Agent* (Agente), o qual é descrito como um ser que fisicamente controla o movimento do Tema. Se mapearmos os Elementos de *Frame* dos dois *frames*, notaremos que o papel abstrato relativo ao Remetente é o de Agente, presente no *frame* de *Bringing*. Assim, o FrameGrapher se faz útil ao revelar as funções macrotemáticas de EFs que foram especificados.

Os Elementos de *Frame* se repartem em nucleares, periféricos e extra-temáticos. Os primeiros instanciam um componente conceitualmente necessário ao *frame*. Existem algumas propriedades que auxiliam na definição do que seja um elemento de *frame* nuclear, a saber: quando um EF precisa sempre ser explicitamente especificado ou sua omissão recebe uma interpretação definida e quando a semântica de um EF não pode ser prevista a partir de sua forma. Já os EFs periféricos são aqueles que introduzem

informação adicional, independente ou distinta em relação ao evento principal. Esse tipo de EF não caracteriza um *frame* (RUPPENHOFER ET AL., 2010). Valendo-nos novamente do *frame* de Transação_Comercial, descrito no Quadro 2, os elementos *Buyer* (Comprador), *Goods* (Mercadorias), *Money* (Dinheiro) e *Seller* (Vendedor) são nucleares, uma vez que são necessários para que se evoque o *frame*; *Means* (Meio), *Rate* (Taxa) e *Unit* (Unidade) são elementos periféricos, não responsáveis por instanciar esse *frame*.

Existem, ainda, os EFs extra-temáticos, os quais podem figurar em um dado *frame*, porém, não pertencem a tal cena. Geralmente, as palavras que são etiquetadas como instanciando um EF extra-temático são ULs em um *frame* distinto. Utilizando-nos mais uma vez do *frame* de Transação_Comercial, considere a sentença (5):

(5) “[Alice Comprador] **comprou** o [caderno Mercadorias] [ontem à tarde Tempo]”.

Nesse exemplo, o EF “ontem à tarde”, o qual refere-se ao Tempo em que a compra fora realizada, é extra-temático. Assim, uma palavra como “ontem” seria uma UL do *frame* de Tempo.

Quando os Elementos de *Frame* nucleares não vêm expressos na sentença, são anotados como Instanciações Nulas. Nesses casos, há três possibilidades: a Instanciação Nula Definida (IND), a Instanciação Nula Indefinida (INI) e a Instanciação Nula Construcional (INC). Observem-se os seguintes exemplos, a fim de se elucidarem os diferentes tipos de Instanciações Nulas:

(6) [O garoto Agente] **chutou** [para o gol Alvo]. [Tema = IND]

(7) [A moça Comedor] não quis **comer**. [Comestíveis = INI]

Em ambas as sentenças, o EF nuclear que ocupa a posição de Objeto não foi exposto. Quando o EF omitido faz referência a uma entidade em particular, temos uma IND, caso da sentença (6), na qual é evidente que o garoto chutou uma bola, isto é, o EF omitido refere-se à entidade “bola”. Por outro lado, quando a referência é feita a uma categoria, trata-se de uma INI, o exemplo (7), no qual a moça não quis comer alguma comida, ou seja, a referência é feita à categoria “comida”.

Por fim, veja o exemplo a seguir:

(8) [O vestido _{Mercadoria}] foi **comprado** ontem. [Comprador = INC]

Em (8), o Comprador, elemento nuclear, não é expreso. Como essa supressão é constitutiva da construção passiva, a chamamos de INC.

Além da anotação detalhada contendo exemplos de sentenças, definição dos *frames* e seus elementos, a FrameNet descreve as relações semânticas existentes entre os *frames*. Ruppenhofer et al. (2010) argumentam que a definição de tais relações traz benefícios, como, por exemplo, o refinamento na compreensão acerca de um *frame* demasiado complexo relacionando-o a um *frame* mais concreto.

As relações são assimétricas, de modo que o *frame* mais complexo pode ser chamado de Superframe, e o menos complexo, Subframe. Em cada tipo de relação, o Superframe e o Subframe receberão nomes específicos relacionados ao tipo de relação em questão. A seguir, apresentaremos as relações presentes na FrameNet.

Conforme a descrição encontrada no *Book* (RUPPENHOFER, et al., 2010), na relação de **Herança**, o que for verdadeiro para o *frame* mãe deve corresponder a algo mais específico no *frame* filho – por exemplo, os EFs presentes no *frame* mãe devem corresponder de modo mais específico aos EFs do *frame* filho. O EF Agente no *frame* de Causar_Mover é aquele cuja ação causa o movimento de um Tema. O *frame* de Passar herda de Causar_Mover; assim, de modo mais específico, no *frame* filho, o Agente é aquele que passa o Tema, como em “Ana me passou a bola”, em que Ana é o Agente e bola é o Tema. Em Passar, o movimento realizado pelo Agente é especificado em relação ao Causar_Mover. Ademais, em uma relação de Herança, caso o *frame* mãe tenha subframes que especificam subeventos nele contidos, o *frame* filho deve conter também subeventos mais específicos que, de alguma forma, espelhem o que é proposto para o *frame* mãe de maneira menos genérica.

Alguns *frames* são complexos na medida em que se referem a sequências de estados e transições, cada um dos quais podendo ser separadamente descrito como um *frame*. Esses *frames* separados, os

subframes, são relacionados aos *frames* complexos através da relação de **Subframe**. Os EFs do *frame* complexo podem ser identificados com os EFs das subpartes, embora nem todos necessitem ser relacionados. O *frame* de Comércio_transferência_de_dinheiro é um subframe de Transação_Comercial – *frame* mais complexo. No Subframe, o foco está na transferência de dinheiro, isto é, uma das microcenas de Transação_Comercial, que envolveria, ainda, a transferência da mercadoria trocada pelo dinheiro.

A relação de **Precedência** ocorre somente entre dois *subframes* componentes de um único *frame* complexo. Essa relação captura a ordenação temporal de subeventos dentro de um evento complexo (PETRUCK & MELO, 2012). O *frame* de Dormir, por exemplo, precede o *frame* de Acordar que, por sua vez, precede o *frame* de Estar_acordado. Os três *frames* – entre outros – pertencem ao *frame* de Ciclo_de_dormir_e_acordar – evento complexo.

A relação de **Uso** designa um *frame* filho que pressupõe o *frame* mãe como *background*. Essa relação é usada quase exclusivamente para casos em que uma parte da cena evocada pelo *frame* filho refere-se ao *frame* mãe. O *frame* de Deixado_por_fazer usa o *frame* de Finalidade, na medida em que se pressupõe no primeiro que, ao se deixar algo por fazer, existe a finalidade de fazê-lo em algum momento.

A relação de **Perspectiva** indica a presença de, pelo menos, dois pontos de vista que podem ser assumidos pelo *frame* neutro. Geralmente, o *frame* neutro é não-lexical e não-perspectivizado. No âmbito da Transação_comercial, os *frames* de Comércio_pagar e Comércio_coletar se relacionam com o *frame* de Comércio_transferência_de_dinheiro – *frame* neutro, não-lexical – pela relação de Perspectiva, a saber, a do Comprador e a do Vendedor respectivamente.

Os *frames* estativos, incoativos e causativos se relacionam através das relações de **Causativo de** e **Incoativo de**. Sentenças como “Ana e Paula secaram o pano” evocam o *frame* Causar_estar_seco, o qual é causativo do *frame* de Estar_seco, exemplificado em “O pano estava seco”. Por outro lado, sentenças como “O pano secou” evocam o *frame* Tornar-se_seco, incoativo de Estar_seco.

Por fim, em casos em que há grupos de *frames* que possuem semelhanças e que, por essa razão, deveriam ser comparados e contrastados,

cada um dos *frames* em questão têm uma relação de **Veja-também**¹¹ com um membro representativo do grupo.

Todas as relações acima descritas podem ser visualizadas no FrameGrapher, mencionado anteriormente. Na Figura 7, outra imagem da ferramenta é disponibilizada, na qual diversos *frames* aparecem relacionados ao *frame* de *Commercial_transaction*. As setas em cores diferentes indicam os variados tipos de relação existentes entre os *frames*.

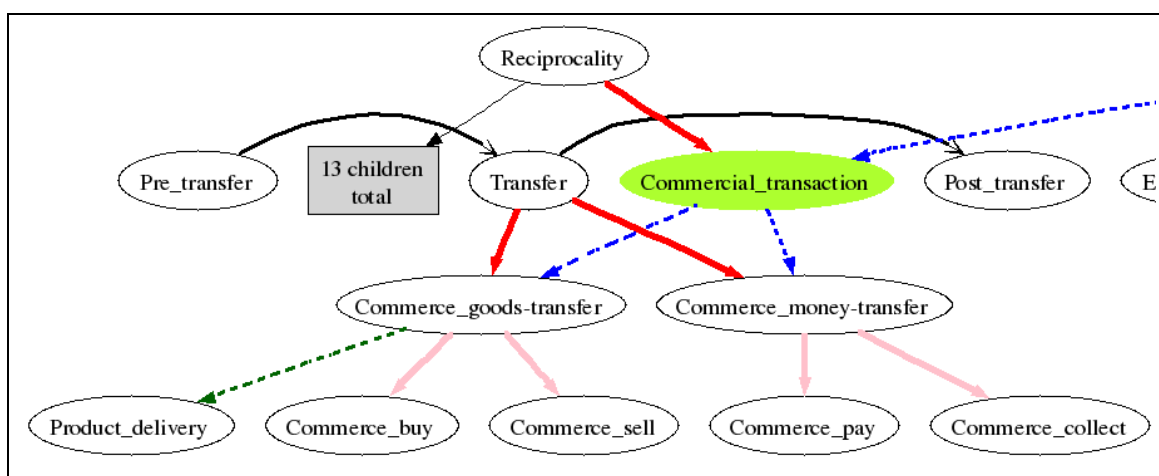


Figura 7: FrameGrapher: Commercial_transaction

A seta azul, por exemplo, representa a relação de Subframe, o que significa que os *frames* de *Commerce_goods-transfer* e *Commerce_money-transfer* são subframes de *Commercial_transaction*. A seta vermelha, por sua vez, indica relações de herança. Assim, *Commerce_goods-transfer* e *Commerce_money-transfer* herdam de *Transfer*.

Ao fim de cada anotação, realizada em camadas conforme descrito no capítulo de Metodologia, são feitas tabelas de sumariamento, as quais evidenciam os padrões de valência sintático-semânticas das ULs de cada *frame*. Através dessas tabelas, é possível visualizar as possibilidades combinatórias que essas ULs apresentam.

A noção de valência é introduzida por Fillmore em trabalhos anteriores (cf. FILLMORE, 1970 *inter alia*). Vejamos os exemplos a seguir:

¹¹ Essa não é propriamente uma relação semântica; trata-se de uma relação proposta pelos lexicógrafos da FrameNet como forma de guiar o consulente da base de dados.

- (9) [João_{Agente}] bateu [na janela_{Paciente}]
 (10) [João_{Agente}] bateu [na janela_{Paciente}] [com um martelo_{instrumento}]

Observando as sentenças (9) e (10), percebe-se que, em português, é possível ter ou não um instrumento. Nesse tipo de análise já se apresenta a ideia de valência: entidades sintáticas em uma frase, as quais preenchem papéis semânticos. As possibilidades de combinação indicam frases gramaticais em uma dada língua.

Vejamos no Quadro 3 um exemplo de tabelas de sumariamento dos padrões de valência encontrados no *frame* de Commerce_buy com a UL *buy* (comprar).

Frame Element	Number Annotated	Realization(s)
Buyer	(98)	CNI.-- (25) DNI.-- (1) NP.Ext (68) PP[by].Dep (4)
Goods	(98)	DNI.-- (13) NP.Ext (13) NP.Obj (63) Sinterrog.Dep (2) AJP.Dep (2) NP.Appositive (1) NP.Dep (5)
Manner	(3)	AVP.Dep (2) PP[in].Dep (1)
Means	(1)	PPing[by].Dep (1)
Money	(25)	PP[at].Dep (9) PP[with].Dep (5) PP[for].Dep (9) NP.Ext (2)
Place	(1)	PP[at].Dep (1)

Purpose	(1)	PP[for].Dep (1)
Recipient	(15)	PPing[for].Dep (2) PP[for].Dep (8) NP.Obj (4) PP[as].Dep (2)
Seller	(18)	PP[from].Dep (17) PP[over].Dep (1)
Time	(3)	Sub.Dep (1) AVP.Dep (2)

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns			
<u>1</u> TOTAL	Buyer	Buyer	Goods	
(1)	NP Ext	NP Ext	NP Obj	
<u>1</u> TOTAL	Buyer	Buyer	Goods	Money
(1)	NP Ext	NP Ext	DNI --	PP[at] Dep

Quadro 3: Sumariamento dos padrões de valência

Na primeira tabela é possível ver os EFs da UL *buy*, a quantidade de vezes em que aparecem nas anotações das sentenças, e suas realizações, isto é, o tipo sintagmático no qual o EF aparece ou, se esse for omitido na sentença, o tipo de instanciação nula em que se manifesta.

A segunda tabela¹² apresenta cada padrão encontrado para a UL *buy*, ou seja, cada possibilidade combinatória de seus EFs em termos de sua realização sintática. Assim, no primeiro padrão visualizamos dois EFs *Buyer* (Comprador) que aparecem como um sintagma nominal e com a função gramatical de Externo; o outro EF é *Goods* (Mercadoria), que também se apresenta em um sintagma nominal, mas tem função gramatical de Objeto. Já o segundo padrão varia nos seguintes pontos: para o EF *Goods* temos uma instanciação nula definida (IND) e um novo EF Figura na sentença, *Money*

¹² Optou-se por exibir somente os dois primeiros padrões encontrados para a UL em questão.

(Dinheiro), o qual apresenta-se em um sintagma preposicionado e tem função gramatical de Dependente.

Conforme mencionado anteriormente, essas tabelas emergem das análises em camadas das ULs. Observemos na Figura 8 um exemplo de como é realizada a anotação em camadas:

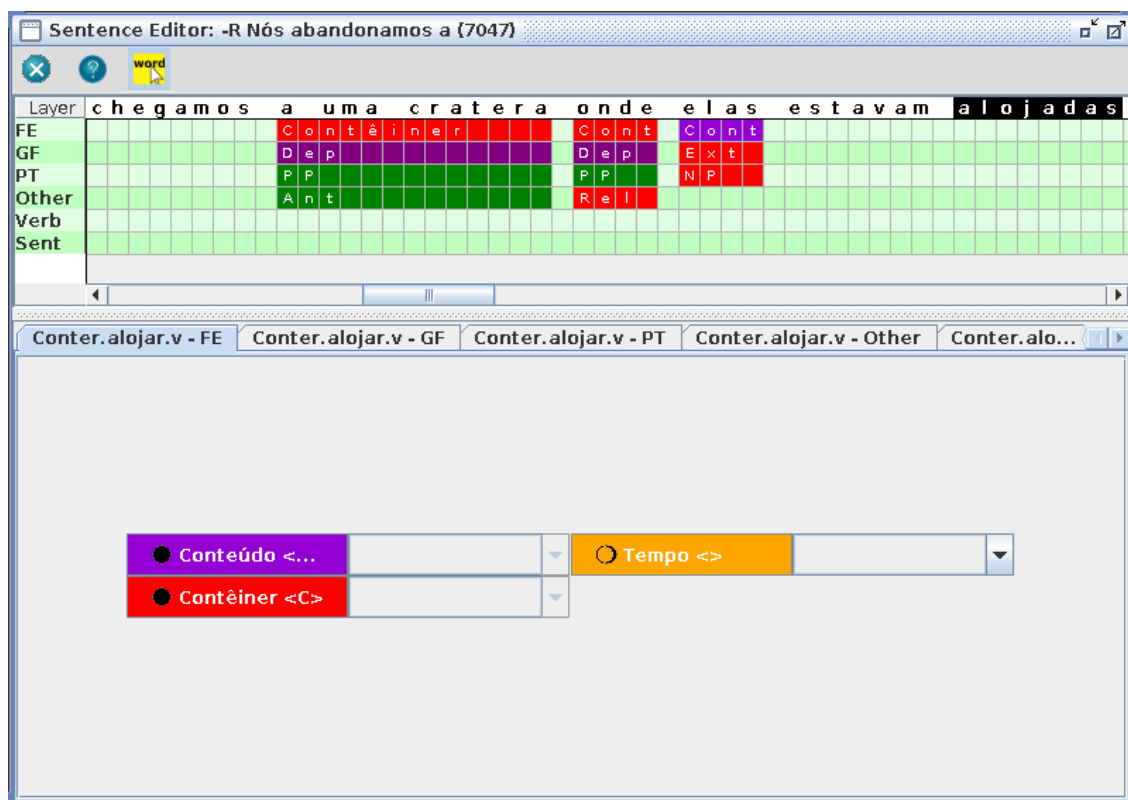


Figura 8: Anotação em camadas da UL Alojjar

A anotação presente na Figura 8 é referente à UL Alojjar, do *frame* de Conter. Em tal anotação, observamos na camada EF (*FE*) os EFs nucleares, Contêiner “a uma cratera”, recuperado pelo relativo “onde” – sinalizado na camada Outros (*Other*) como “Relativo” –, e Conteúdo “elas”. Na camada FG (*GF*), vemos que Contêiner é um Dep, ao passo que o Conteúdo é o Ext. O TS (*PT*) é assinalado como SP para o Contêiner e SN para o Conteúdo.

Ao adotar a anotação em camadas e dispor elementos de *frame* sobre função gramatical e tipo sintagmático, a FrameNet demonstra a base construcionista que permeia sua perspectiva teórico-metodológica. Elementos de *Frame* são a contraparte semântica na análise, ao passo que a função gramatical (FG) e o tipo sintagmático (TS) são sua contraparte formal. Sabendo

que construções são o pareamento de uma forma a um significado, a disposição em camadas revela a inspiração construcionista da análise lexicográfica levada a cabo na FrameNet, uma vez que em tais tabelas visualizam-se padrões de combinação entre **forma** (FG e TS) e **sentido** (EF).

Desse modo, partimos, agora, para a discussão da Gramática das Construções contemplando seus vários matizes de modo a estabelecer as bases sobre as quais discutiremos em que medida a implantação de um repertório de construções – Constructicon – se faz necessária, dado o viés construcionista das análises já disponibilizadas pela anotação lexicográfica.

2.2 A Gramática das Construções

Nesta seção, abordaremos duas vertentes da Gramática das Construções, a saber, a Gramática Cognitivista das Construções, desenvolvida por Goldberg, e a Gramática das Construções baseada em Unificação, proposta por Fillmore e Kay. Ambas serão apresentadas de modo a apontar suas características teórico-analíticas, objetivos e suas principais diferenças.

A escolha dos dois modelos acima referidos – existem outros modelos em Gramática das Construções – não é sem razão. No que tange à Gramática Cognitivista das Construções, a escolha foi feita uma vez que a família de construções escolhida para implantarmos o Constructicon, a *Para Infinitivo*, foi anteriormente analisada por Torrent (2009) seguindo-se os princípios desse modelo. Quanto ao segundo modelo, Gramática das Construções baseada em Unificação, a escolha se deve ao fato de que o Constructicon a ser apresentado na seção seguinte desenvolveu-se com base nessa proposta.

Antes de aprofundar nossa investigação acerca desses dois modelos, é importante esclarecer alguns dos princípios comuns a todas as vertentes em Gramática das Construções. O primeiro deles é a postulação de que qualquer tipo de expressão linguística, desde as mais simples às mais complexas, pode ser considerado uma unidade composta de uma correspondência entre forma e significado ou função¹³. A proposta de uma gramática baseada em construções

¹³ É importante observar que Fillmore admite a existência de algumas construções meramente formais – tais como a construção da inversão do sujeito e auxiliar do inglês. Goldberg, por outro lado, propõe que toda construção tenha significado ou função.

surge para abarcar os casos em que uma abordagem centrada em regras não é suficiente. Inicialmente, Fillmore e colaboradores (1988) propõem tal abordagem com o intuito de explicar as ocorrências mais irregulares da língua. Se tais princípios fossem capazes de explicar os casos periféricos, também explicariam as regularidades.

Pressupondo-se o pareamento entre forma e significado / função nos níveis lexical, morfológico e sintático – ou mesmo discursivo –, assume-se uma continuidade entre léxico e sintaxe. Esta é uma hipótese assumida pela Linguística Cognitiva na medida em que esta não concebe léxico e sintaxe como blocos rigidamente separados. Assim, itens lexicais, como *flor*, expressões idiomáticas, como *dar mole*; ou padrões de estrutura argumental, como $[SN_{Agente} V SN_{Paciente}]$ são igualmente construções.

2.2.1 A Gramática Cognitivista das Construções

Um princípio fundamental para a Gramática Cognitivista das Construções é o de que construções, isto é, correspondências entre forma e significado, são as unidades básicas de uma língua. Inicialmente, Goldberg atribui o termo construção a qualquer item linguístico cujo significado e/ou forma não seja composicionalmente derivado de outras construções existentes na língua (1995, p.4).

Posteriormente, a noção de construção é expandida, sendo definida da seguinte forma:

Qualquer padrão linguístico é reconhecido como uma construção desde que algum aspecto de sua forma ou função não seja estritamente previsível a partir das partes que o compõem ou a partir de outras construções reconhecidamente existentes. Além disso, padrões são armazenados como construções, mesmo os totalmente previsíveis, desde que ocorram de forma suficientemente frequente. (GOLDBERG, 2006, p.5)¹⁴

Através dessa reformulação no conceito de construção, Goldberg amplia seu escopo analítico ao adotar como unidades de análise não só as construções

¹⁴ Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions cognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.

mais idiomatizadas. Padrões tratáveis composicionalmente também são considerados construções, desde que sejam frequentes na língua.

Por consequência dessa concepção, não há uma divisão estrita entre léxico e sintaxe; em vez disso, os dois campos são compreendidos como um *continuum*. Goldberg (2006) argumenta que construções lexicais e construções sintáticas se diferenciam quanto à sua complexidade interna, mas ambas são essencialmente o mesmo tipo de estrutura, um pareamento de forma e significado. Negar uma divisão estrita entre esses dois itens não significa negar a existência de qualquer restrição distintamente morfológica ou sintática. O que se reivindica é que existem semelhanças básicas entre os dois tipos de construção.

Um dos interesses dos pesquisadores que seguem essa vertente é descrever as condições sob as quais uma construção pode ser empregada de modo apropriado, uma vez que isso é parte da competência do falante. Disso decorre a convicção de que fatores semânticos e pragmáticos são cruciais para se compreenderem as restrições em construções gramaticais. Vale lembrar que essa vertente rejeita uma divisão estrita entre Semântica e Pragmática.

Finalmente, é importante ressaltar que a abordagem construcionista deixa de reconhecer apenas os itens lexicais como elementos indicadores de sentido. Nesse viés, as construções, dentre elas as de estrutura argumental, também podem contribuir com o significado. Esse ponto de vista será discutido na próxima subseção.

2.2.1.1 Significados Verbais e Significados Construcionais

Adele Goldberg (1995; 2010) trava uma discussão a respeito da distinção entre os significados evocados por verbos e aqueles evocados por construções, o que revela o lugar de destaque dado por essa abordagem à relação entre tais elementos.

A autora explica que um conhecimento rico em termos de Semântica de *Frames* acerca de verbos é necessário para, por exemplo, o uso adequado de advérbios e adjuntos, ou no processo de preempção, entendido como o bloqueio do uso de determinado item em um padrão paradigmático mais

genérico. No caso dos advérbios, a autora exemplifica a situação com as seguintes sentenças (GOLDBERG, 1995, p.29 e 30):

- (11) Joe walked into the room slowly.
 (12) ??Joe careened into the room slowly.

A segunda sentença é problemática visto que o verbo usado, *careened*, implica um movimento rápido e descontrolado, o que torna estranha a presença do advérbio *slowly* – lentamente. Assim, conhecer os aspectos sintáticos relevantes não é suficiente para determinar o uso de um advérbio em uma sentença.

Em processos de preempção, o conhecimento do *frame* concernente ao verbo é também fundamental. Goldberg cita os casos de supergeneralização que crianças fazem na fase de aquisição da linguagem. É comum regularizar verbos irregulares em situações como “eu fazi” em vez de “eu fiz”. A preempção ocorre quando a criança, conhecendo o termo sinônimo “eu fiz” passa a bloquear o termo regular, mas, para que isso aconteça, é preciso que a as duas formas tenham semântica idêntica. Portanto, o conhecimento sobre o *frame* evocado por um verbo precisa ser conhecido e reconhecido pelo falante para que a preempção ocorra.

Em relação ao significado construcional, Goldberg pontua que, ao invés de se postularem regras lexicais, emprega-se uma análise baseada na polissemia da construção: “a mesma forma é pareada a sentidos diferentes, mas relacionados” (GOLDBERG, 1995, p.33). Ela sugere que se admita um sentido central específico da construção e que se postulem sentidos separados, mas relacionados, referentes a classes de verbos específicas.

Sua sugestão vai de encontro à de se postular um único sentido abstrato para a construção e permitir que a semântica do verbo preencha o significado, visto que, de acordo com a autora, essa proposta falha em explicar os dados. Goldberg afirma que o significado uniforme reivindicado para a construção ditransitiva, por exemplo, é questionável em vários aspectos (GOLDBERG, 1995, p. 34 e 35). Além disso, uma abordagem abstracionista não consegue apreender que a noção de transferência, em geral, é básica a essa construção¹⁵, o que torna alguns verbos mais prototípicos – tais como *dar*.

¹⁵ Goldberg faz um experimento em que pergunta a algumas pessoas sem formação em Linguística o significado da palavra “*topamased*” na sentença “*She topamased him something.*”

Segundo a autora, uma abordagem polissêmica reconhece o papel central desempenhado pelo significado da construção.

Após argumentar que a construção ditransitiva tem como significado central “transferência bem sucedida”, Goldberg afirma que construções de nível oracional básico podem ser vistas como designadoras de cenas relevantes da vida humana e, então, propõe a Hipótese da Codificação das Cenas, em que “construções que correspondem a tipos básicos de sentença codificam como seus significados centrais, tipos de eventos que são básicos à experiência humana”¹⁶ (GOLDBERG, 1995, p.39).

A autora discute que determinadas combinações de papéis que designam cenas relevantes da vida humana são associadas a construções de estrutura argumental, as quais dividem o mundo em tipos de evento discretamente classificados. Por outro lado, os verbos são relacionados a significados ricos em termos da Semântica de *Frames*.

Muito se especula acerca das restrições impostas à combinação de eventos que um único verbo pode designar. Croft (1991; 2005) sustenta que verbos designam somente eventos simples, ou eventos cujos subeventos são causalmente relacionados. Entretanto, Goldberg apresenta contra-argumentos à tese de Croft, como, por exemplo, o caso do verbo “*braised*”, o qual designa dois subeventos, a saber, cozinhar e assar – em que um evento não está completamente dentro da extensão temporal do outro –, não relacionados causalmente – um evento não é suficiente para acarretar o outro (GOLDBERG, 2010). Portanto, verbos nem sempre indicam eventos relacionados causalmente.

Levin e Rappaport Hovav (apud Goldberg 2010) propõem outra restrição: a de que um único verbo não pode designar, ao mesmo tempo, modo e resultado, ou modo e mudança de lugar. Novamente, Goldberg mostra que o verbo “*climb*” contradiz a tese dos autores, uma vez que implica modo – escalada – bem como movimento direcionado – para cima.

Dentre dez pessoas, seis responderam “*give*”, o que aponta para uma relação estreita entre a construção ditransitiva e o verbo em questão.

¹⁶ *Scene Encoding Hypothesis*: Constructions which correspond to basic sentence types encode as their central senses event types that are basic to human experience.

Assim, Goldberg propõe a restrição que rege a relação entre eventos designados por um único verbo: a restrição do *Frame* Convencional.

Para uma situação ser rotulada por um verbo, a situação ou experiência pode ser hipotética ou histórica e não precisa ser experienciada diretamente, mas é necessário que a situação ou experiência evoque uma unidade cultural que é familiar e relevante para aqueles que usam a palavra¹⁷. (GOLDBERG, 2010)

Por essa definição, entende-se que a única restrição é a de que o significado de um verbo não pode conter dois ou mais subeventos não relacionados por um *frame* semântico. Todo verbo designa sempre um *frame* já estabelecido, reconhecido como uma unidade cultural. Para que um falante mencione o verbo “matar”, ele não precisa ter vivenciado essa experiência; entretanto, é necessário que o termo seja familiar a ele e a quem ele se dirige.

Goldberg explica, ainda, que um *frame* semântico oferece as condições para a existência de uma instanciação lexical. No entanto, a existência de um *frame* não acarreta a existência de um verbo que o evoque, isto é, a existência de uma cena comum aos falantes não garante a existência de um verbo que rotule essa cena.

O que se observa no caso das construções, porém, é distinto do que se passa com os verbos. Sentenças de tipo simples são diretamente correlacionadas a estruturas semânticas. Por exemplo, a sintaxe da construção ditransitiva relaciona-se à seguinte estrutura semântica:

[(Suj) V Obj1 Obj2] \implies X CAUSA Y a RECEBER Z

Essa construção de estrutura argumental designa um *frame* semântico geral e abstrato. O significado construcional, assinala Goldberg, emerge a partir de uma generalização de muitos verbos diferentes, focalizando, assim, o *frame* perfilado, e não o *frame* de *background* (2010).

Dessa forma, construções e verbos podem veicular significados separadamente. Se tais elementos podem designar eventos distintos, é preciso investigar se há alguma restrição semântica geral para realizar a combinação entre eles.

¹⁷ For a situation to be labeled by a verb, the situation or experience may be hypothetical or historical and need not be directly experienced, but it is necessary that the situation or experience evoke a cultural unit that is familiar and relevant to those who use the word.

Goldberg (2010) apresenta ocorrências em que os eventos indicados pelo verbo e pela construção são vinculados causalmente especificando-se os meios, o resultado ou o instrumento envolvido em algum ato. Há, ainda, ocorrências em que os eventos não se relacionam causalmente. Nesses casos, a construção permite que o verbo indique uma condição para a realização do ato designado pela construção.

Visto isso, a seguinte questão se impõe: a restrição do *Frame* Convencional incidiria, também, sobre combinações entre verbos e construções? Em outras palavras, a associação de verbos e construções sempre evocará *frames* já estabelecidos, como no caso dos verbos? Goldberg argumenta que não é necessário haver *frames* pré-existentes que envolvam a semântica de um verbo e uma construção de estrutura argumental, exemplificando seu argumento com a sentença em (13):

(13) I *microwaved* her some leftover noodles.

A autora diz que seria um exagero afirmar que existe um *frame* que envolva a cena de esquentar algo no microondas para uma pessoa. Portanto, embora verbos e construções frequentemente se combinem designando *frames* pré-existentes, podem se combinar gerando sentidos completamente novos. Esse novo dado caracteriza uma diferença importante entre verbos e associações de verbos e construções.

2.2.1.2 Motivação e Relações de Herança entre Construções

Goldberg (1995) explica que construções compõem um conjunto estruturado de informação inter-relacionada. Hierarquias de herança e motivação são utilizadas para descrever a forma como as construções se relacionam entre si, de modo a revelar o conhecimento que temos acerca da língua. Dessa forma, a Gramática Cognitivista das Construções vale-se de quatro princípios organizadores da língua.

O primeiro deles refere-se à questão da motivação e é definido da seguinte forma:

Princípio da Motivação Maximizada: Se a construção A é sintaticamente relacionada à construção B, então, o sistema da construção A é motivado

na medida em que é relacionado semanticamente à construção B (cf. Haiman 1985a; Lakoff 1987). Tal motivação é maximizada.¹⁸ (GOLDBERG, 1995, p.67)

A motivação é um ponto central para a Gramática Cognitivista das Construções, a qual busca fornecer tal componente para cada construção postulada (GOLDBERG, 2006, p.217). Goldberg afirma que a motivação é empregada para se explicar porque é possível e natural que um dado pareamento de forma e significado exista numa dada língua.

Lakoff (1987) pontua que o objetivo em se motivar cada construção é uma tentativa de tornar a teoria adequada em sua explicação, e propõe uma definição para o termo “motivação”, afirmando que uma dada construção é motivada na medida em que sua estrutura é herdada de outras construções na língua.

Goldberg esclarece, ainda, que motivação diferencia-se de previsão, na medida em que reconhecer a motivação para uma construção não acarreta na existência dessa construção naquela ou em outra língua. Ao mesmo tempo, não se pode negar que muitas instâncias na língua não sejam absolutamente arbitrárias – ainda que não se possa prever que o algarismo 19 receberá o nome *dix-neuf*, em francês, não se pode negar que exista uma motivação por trás dessa designação (GOLDBERG, 1995, p.69). Nesse sentido, a motivação se posicionaria entre a previsibilidade e a arbitrariedade.

A motivação se apresenta como índice norteador de análises. É o que podemos observar nos resultados dos estudos de Gelman, Wilcox e Clark (1989 *apud* GOLDBERG, 1995), em que crianças se mostraram mais propensas a aprender nomes de termos subordinados, quando esses termos se somavam a outros de nível básico que elas já conheciam. Por exemplo, ao aprenderem um nome para um tipo específico de caminhão, crianças teriam mais facilidade em aprender algo como *caminhão-puca* do que simplesmente *puca*. Embora tal resultado não fosse esperado, uma vez que a estrutura do nome composto é mais elaborada, a motivação surge como explicação natural ao fenômeno:

“Crianças aprendem mais facilmente novos termos para conceitos relacionados a outros que já lhes são familiares, quando os novos termos

¹⁸ The Principle of Maximized Motivation: If construction A is related to construction B syntactically, then the system of construction A is motivated to the degree that it is related to construction B semantically. Such motivation is maximized.

são sistematicamente relacionados aos termos utilizados para os conceitos familiares” (GOLDBERG, 1995, p.71).¹⁹

Um importante preceito recuperado ao se empregar a motivação nas análises linguísticas é o de que elementos num sistema influenciam-se mutuamente, mesmo quando não interagem literalmente, como afirma Goldberg (1995, p. 72). A pesquisadora assinala, ainda, que diversos fenômenos fonológicos apontam para o fato de que os falantes buscam regularidades e padrões; e quando esses ainda não estão disponíveis, os falantes tendem a conferir tais elementos à língua.

Outros três princípios se seguem ao de Motivação, quais sejam, Princípio da Não Sinonímia, Princípio da Expressividade Máxima, e Princípio da Economia Máxima, cujas definições vemos a seguir:

Princípio da Não-Sinonímia: Se duas construções são sintaticamente distintas, elas devem ser semântica ou pragmaticamente distintas (cf. Bolinger, 1968; Haiman, 1985a; Clark, 1987; MacWhinney, 1989).

Princípio da Expressividade Máxima: O inventário das construções é maximizado para fins comunicativos.

Princípio da Economia Máxima: O número de construções distintas é minimizado tanto quanto possível, tendo em conta o III Princípio (Haiman, 1985a). (GOLDBERG, 1995, p.67 e 68).²⁰

Haiman (1985a *apud* GOLDBERG, 1995) usa o princípio do isomorfismo para explicar o funcionamento das línguas, comparando a representação que um mapa faz da condição geográfica de um dado lugar, à representação feita pela língua acerca da realidade. O pesquisador observa que tal princípio cobre dois aspectos da língua: se por um lado, diferenças na forma acarretam diferenças no significado – o que aponta para o Princípio da Não-Sinonímia –,

¹⁹ Children learn new terms for concepts which are related to other, already familiar concepts more easily when the new terms are systematically related to the terms for the familiar concepts.

²⁰ *The Principle of Non Synonymy*: If two constructions are syntactically distinct, they must be semantically or pragmatically distinct (cf. Bolinger 1968; Haiman 1985a; Clark 1987; MacWhinney 1989). Pragmatic aspects of constructions involve particulars of information structure, including topic and focus, and additionally stylistic aspects of the construction such as register (...).

The Principle of Maximized Expressive Power: The inventory of constructions is maximized for communicative purposes.

The Principle of Maximized Economy: The number of distinct constructions is maximized as much as possible, given Principle III (Haiman 1985a).

diferenças no significado deveriam levar a diferenças na forma – Princípio da Expressividade Máxima.

Haiman explica, ainda, que desvios dessa regra ocorrem em casos de polissemia e homonímia. Todavia, tais exceções podem ser atribuídas a uma necessidade de simplificação – o que o Princípio da Economia Máxima abarca. Desse modo, o Princípio da Economia Máxima e o Princípio da Expressividade Máxima restringem-se mutuamente, na medida em que o primeiro reduz a abundância de construções, e o último constitui um sistema que proporciona um rótulo distinto para cada item distinto (GOLDBERG, 1995).

Relações de herança vêm sendo usadas como um método para capturar generalizações linguísticas. Relações de herança assimétricas são postuladas entre construções relacionadas semântica e sintaticamente, a fim de se estabelecer relações de motivação. Assim, uma construção A motiva a construção B se, e somente se, B herda de A.

A Gramática Cognitivista das Construções emprega as relações de herança sob três aspectos a serem destacados a seguir, quais sejam, herança múltipla, modo normal de herança e cópia real. Nessa abordagem, uma construção em uma hierarquia pode herdar de mais de uma construção, isto é, a herança múltipla é permitida. Além disso, o modo normal de herança, o qual apreende generalizações parciais, é também usado, o que significa que subregularidades e exceções são admitidas. Por fim, trabalhar com cópia real implica que as construções dominadas contenham toda a informação que as construções dominantes possuem. Desse modo, “cada construção é totalmente especificada, mas é redundante na medida em que a informação é herdada de (por exemplo, partilhada com) construções dominantes” (GOLDBERG, 1995, p.74)²¹.

Nesse sistema, os links de herança são tratados como objetos. Logo, assim como as construções, assume-se que tenham estrutura interna e que sejam hierarquicamente relacionados. Há quatro tipos principais de links de herança: links de Polissemia, links de Subparte, links de Instanciação e links de Extensão Metafórica.

²¹ (...) each construction is fully specified, but is redundant to the degree that information is inherited from (i.e., shared with) dominating constructions.

Nas palavras de Boas, o **link de Polissemia** representa “relações entre subtipos de construções que exibem as mesmas especificações sintáticas, mas diferenciam-se em sua semântica”²² (2010, p.16). Goldberg explica a função dos links de Polissemia:

Os links de polissemia capturam a natureza das relações semânticas entre um sentido particular de uma construção e quaisquer extensões a partir desse sentido. As especificações sintáticas do sentido central são herdadas pelas extensões; por esse motivo, não precisamos estabelecer a realização sintática para cada construção – tais especificações são herdadas da construção dominante.²³ (GOLDBERG, 1995, p.75).

Para esclarecer, vejamos a análise de Goldberg sobre a Construção Ditransitiva. Seu sentido central pode ser representado pelo esquema “X CAUSA Y A RECEBER Z”, exemplificado pela autora na seguinte sentença: “*Joe gave Sally the ball*”. Boas (2010, p.16) explica que, na análise de Goldberg, o sentido central associa-se a certas classes de verbo que denotam ações relacionadas ao ato de dar – tais como *give, pass, hand* –, verbos de causação instantânea de movimento balístico – como *throw* e *toss* –, e verbos de causação contínua em uma direção especificada deitadamente – como *bring* e *take*. A Construção Ditransitiva apresenta, ao total, cinco extensões de sentido, as quais conectam-se, cada uma, ao sentido central por links de Polissemia. Tais links herdam o esquema sintático do protótipo, e as extensões de sentido são associadas a classes de verbo específicas.

Através da quinta extensão da Construção Ditransitiva – por vezes chamada de Construção Benefactiva –, a autora exemplifica de que modo se representam links de Polissemia (Figura 9). A Construção Benefactiva é aquela em que a causação é pretendida e não necessariamente realizada. São exemplos dessa construção sentenças como “*Maria assou-me um bolo*”, na qual, por mais que a intenção de Maria seja a de que eu receba o bolo, isso não necessariamente acontecerá. A informação herdada do sentido central está sinalizada em *itálico* no segundo quadro. O link I_p entre o sentido central

²² (...) relations between subtypes of constructions that exhibit the same syntactic specifications but differ in their semantics.

²³ *Polysemy links* capture the nature of the semantic relations between a particular sense of a construction and any extensions from this sense. The syntactic specifications of the central sense are inherited by the extensions; therefore we do not need to state the syntactic realization for each extension-such specifications are inherited from the dominating construction.

(Construção Ditransitiva) e a extensão (Construção Benefactiva) relaciona causação à causação intencional.

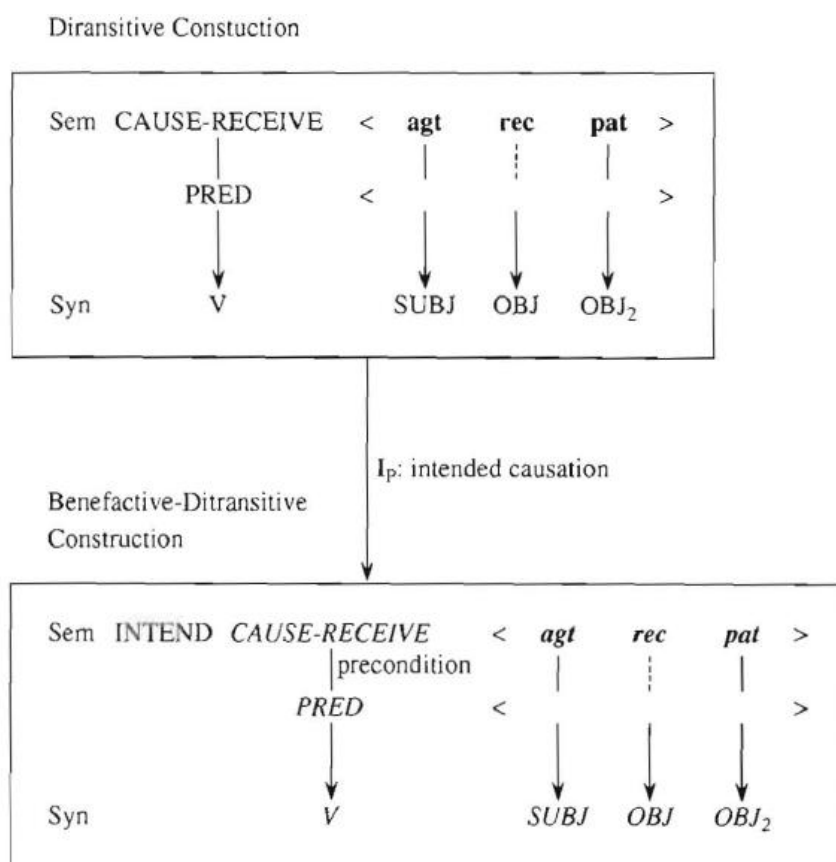


Figura 9 (GOLDBERG, 1995, p.77)

Os **links de Subparte** são usados para identificar relações em que uma construção é uma subparte de outra construção e existe de modo independente. Observe na Figura 10 a diagramação do link de Subparte que a Construção de Movimento Causado tem com relação à Construção de Movimento Intransitivo.

A Construção de Movimento Causado pode ser exemplificada em uma sentença como “Ronaldo rolou a bola pro gol”, em que o movimento da bola foi causado por Ronaldo. A Construção de Movimento Intransitivo, por sua vez, é aquela em que a causa do movimento não é mencionada. Como exemplo, temos a sentença “A bola rolou pro gol”.

As especificações sintáticas e semânticas da Construção de Movimento Intransitivo constituem uma subparte das especificações sintáticas e semânticas da Construção de Movimento Causado. Assim, o Tema que era

objeto na segunda construção, passa a ser sujeito na primeira, já que a Causa da primeira construção não aparece na segunda.

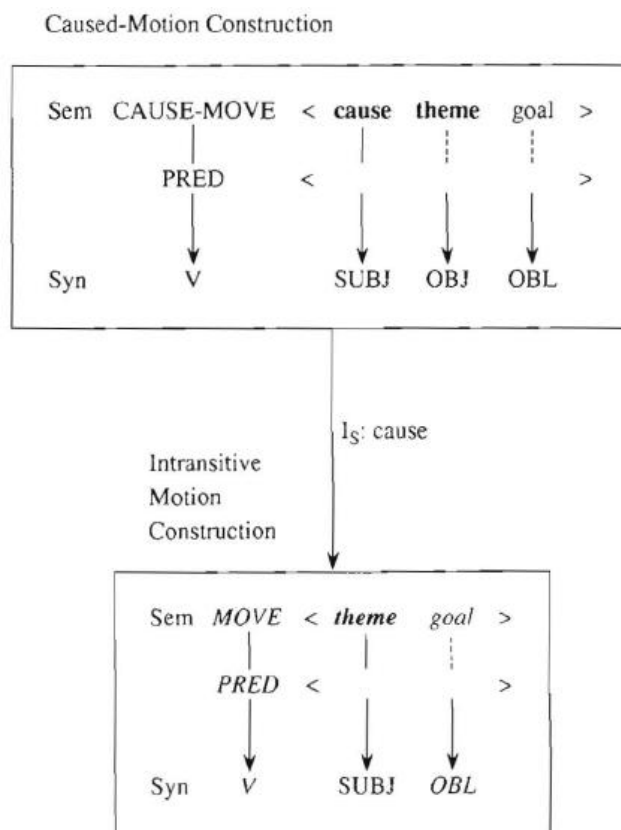


Figura 10 (GOLDBERG, 1995, p.78)

Quanto aos **links de Instanciação**, Goldberg assinala que esses são estabelecidos quando:

uma construção particular é um caso especial de outra construção, isto é, um link de instanciação existe entre construções se, e somente se, uma construção é uma versão mais inteiramente especificada de outra. Itens lexicais particulares que só ocorrem em uma construção particular, são instâncias dessa construção, uma vez que lexicalmente herdam a sintaxe e semântica associada à construção. Estes casos são, portanto, tratados como instâncias da construção parcialmente preenchidas lexicalmente.²⁴ (GOLDBERG, 1995, p.79)

A autora exemplifica com um sentido específico de “*drive*”, como na sentença “*Kim drove Fred crazy*”. Esse sentido ocorre apenas na Construção Resultativa, e é vinculado a essa construção por um link de Instanciação.

²⁴ (...) a particular construction is a *special case* of another construction; that is, an instance link exists between constructions iff one construction is a more fully specified version of the other. Particular lexical items which only occur in a particular construction are instances of that construction, since they lexically inherit the syntax and semantics associated with the construction. These cases are therefore treated as partially lexically filled instances of the construction.

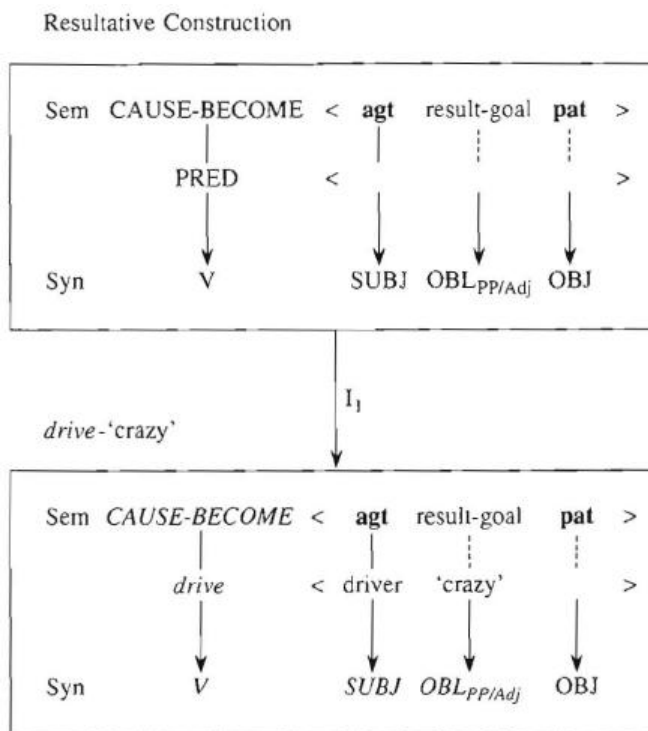


Figura 11 (GOLDBERG, 1995, p.80)

Por fim, construções podem interagir por um **link de Extensão Metafórica**, quando duas construções se relacionam por um mapeamento metafórico. A metáfora especifica o modo como a semântica da construção dominante é mapeada à semântica da construção dominada. Um exemplo é a Construção Habilitativa com *Dar*²⁵, cuja estrutura é herdada da Construção de Transferência de Posse através de um link de Extensão Metafórica, como aponta Salomão (2007b). A representação da Construção Habilitativa com *Dar* pode ser observada na Figura 12:

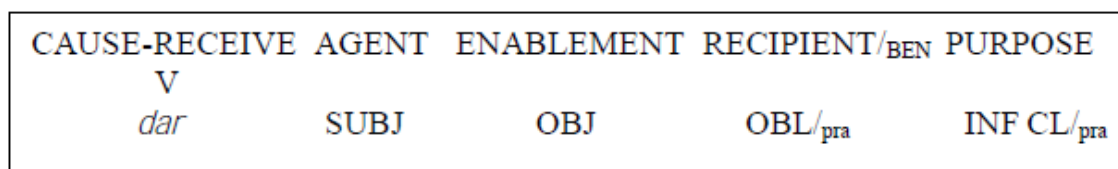
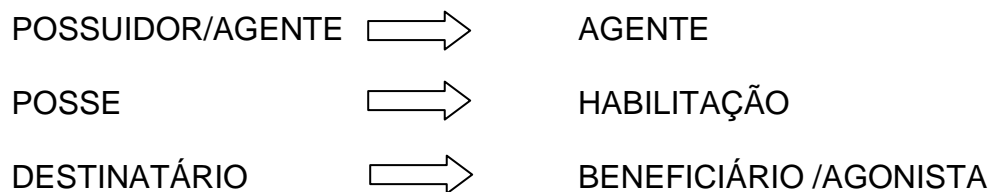


Figura 12 (SALOMÃO, 2007b)

²⁵ A construção habilitativa com *dar*, estudada por Salomão (1990; 2007b), tem sua análise retomada por Torrent (2009; 2011; no prelo), o qual a caracteriza como a Construção de Dativo com Infinitivo. Essa análise será retomada no capítulo 3 desta dissertação.

A metáfora que sustenta a relação entre as duas construções citadas acima é, segundo Salomão (2007b), HABILITAÇÃO É POSSE. Dada essa metáfora, é possível realizar o seguinte mapeamento através desse link:



A linguista explica que a Construção Habilitativa com *Dar* herda a estrutura de *frame* da Construção de Transferência de Posse e se combina com a Construção de Finalidade para compor seu próprio pareamento sintático-semântico (SALOMÃO, 2007b).

Os links de herança entre construções propostos por Goldberg (1995) sustentam as relações de motivação observáveis quando da proposição de redes construcionais. No capítulo 3, mostraremos como esses links foram aplicados por Torrent (2009; 2011; no prelo) na estruturação da família de construções em *Para Infinitivo*.

2.2.2 A Gramática das Construções Baseada em Unificação

De inclinação mais formalista do que a Gramática Cognitivista das Construções, as gramáticas baseadas em unificação (dentre as quais se destaca a *Sign-Based Construction Grammar* – SBCG), buscam incorporar descrições interpretáveis computacionalmente aos pressupostos fundamentais das abordagens construcionistas apresentadas no início deste capítulo.

Sag et al. (2012) afirmam que atualmente existem três comunidades de pesquisa em morfossintaxe: o campo da Gramática Universal – o qual considera as diversas línguas naturais como instanciações de uma única gramática –, o campo Tipológico – voltado a observações descritivas de línguas individuais, em especial, de idiosincrasias e complexidades – e o campo da Gramática Formal (GF), do qual emerge a abordagem a ser tratada nesta seção, a SBCG. A pesquisa no âmbito da GF fundamenta-se matematicamente e busca incorporar práticas computacionais. Ela reúne

trabalhos lexicalistas baseados em restrição, cujas análises são “precisamente formuladas e internamente consistentes” (SAG et al. 2012, p.3). Lee-Goldman (2011) assinala que a SBCG mescla ideias oriundas dos dois primeiros campos supracitados, ampliando o potencial de trabalho na Gramática Universal ao incorporar intuições da Gramática das Construções. Ao mesmo tempo, proporciona à última um maior formalismo à sua base teórica.

Para Sag et al., proporcionar um trabalho formalizado no qual seja possível se aplicarem as ideias e teorias de pesquisadores do campo tipológico pode ser a principal meta da SBCG, uma vez que a formalização presente nessa vertente possibilita “previsões empíricas precisas, aperfeiçoamento de comparabilidade das análises linguísticas e clareza teórica geral” (2012, p.3).

Fillmore e Kay (1999) ressaltam que adotar uma abordagem construcional significa se comprometer com a descrição total de uma língua, isto é, essa gramática deve tratar desde padrões gerais até os mais idiomáticos, através de um sistema explícito de representação, capaz de codificar construções de modo econômico e sem perda de generalização.

A proposta apresentada pelos autores supracitados é, portanto, a de uma gramática explícita, monoestratal e não derivacional, baseada na noção de uma construção gramatical como uma associação convencional da forma linguística ao conteúdo (FILLMORE & KAY, 1999, p.2).

Na SBCG, construções recebem uma definição um pouco diferente daquela proposta por Goldberg (vide seção anterior), uma vez que o pareamento de forma e significado é mediado por signos. A noção de signo aqui utilizada se assemelha à de Saussure (2008 [1916]), uma combinação de forma (significante) e sentido (significado); porém, essa noção é ampliada ao incluir associações entre fonologia, morfologia, sintaxe, semântica e informação contextual, como sinaliza Lee-Goldman (2011). A proposta em questão intitula-se “baseada em signos”, na medida em que busca representar adequadamente todos os signos linguísticos, cuja estrutura básica pode ser observada na Figura 13, a qual representa o nome próprio *Kim*:

[PHON	/kim/	
	SYN	PN	
	SEM	‘the intended person named Kim’]

Figura 13 (SAG et al. 2012, p.15)

Na representação acima, PHON refere-se à contraparte fonológica do signo em questão. SYN e SEM correspondem, respectivamente, à informação categorial – nome próprio – e ao significado que o signo apresenta. Nas descrições realizadas pela SBCG, os atributos dos signos são dispostos em forma de traços, aos quais são associados possíveis valores. Podem chegar a seis os traços presentes em uma descrição, os quais listamos a seguir: fonologia (FON), forma (FORMA), estrutura argumental (EST-ARG), sintaxe (SIN), semântica (SEM) e contexto (CTXT).

Lee-Goldman (2011) explica que o primeiro traço, FON, corresponde a um sintagma fonológico. O traço FORMA registra as características morfológicas do signo. Os possíveis argumentos de um signo são descritos pelo terceiro traço, EST-ARG – o qual é similar à Valência, mas diferencia-se por ser definido apenas por signos lexicais. O quarto traço, SIN, vale-se de objetos-sintáticos, os quais, por sua vez, especificam valores para os subtraços²⁶ Categoria (CAT) – nome, verbo, adjetivo, etc – e Valência (VAL) – representa as possibilidades combinatórias de um signo. O quinto traço, SEM, indica o significado do signo e tem como valores os subtraços INDEX – em que cada signo alude a algum objeto no universo do discurso – e FRAME – responsável por listar os elementos que são evocados pelo signo. Por fim, o traço CTXT é requisitado quando traços do contexto são necessários à compreensão de um signo.

Sampaio (2010) nos apresenta um exemplo da representação de um signo feita pela SBCG na Figura 14, na qual 14 notamos os seis traços supracitados. Primeiramente, a realidade fonológica do signo é exibida em FON, seguida de sua FORMA. O traço EST-ARG não está preenchido visto que o signo sob análise não possui argumentos. A parte sintática é descrita em

²⁶ Optou-se por denominá-los “subtraços” em vez de “traços”, com a finalidade única de evitar que o leitor se confunda com os seis traços principais mencionados anteriormente.

termos dos valores CAT – trata-se de um nome – e VAL, que também não contém nenhuma informação, pelo mesmo motivo de EST-ARG. O traço semântico tem o valor INDEX sinalizado com *i* coindexado a Bruno, isto é, a referência feita por esse signo é o sujeito “Bruno”. O valor FRAME indica que foi evocado o *frame* de nome; que o nome evocado é “Bruno”; e que o indivíduo designado é o sujeito que recebe esse nome. O traço CTXT não foi acionado uma vez que nenhum traço do contexto foi necessário para se compreender esse signo.

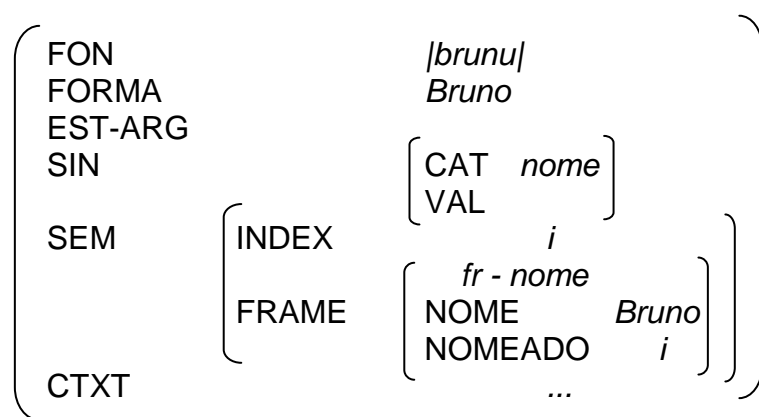


Figura 14: Lexema Bruno (SAMPAIO, 2010, p.36)

Como Goldberg (2006) descreve, a Gramática das Construções Baseada em Unificação é um sistema baseado em traços em que as construções são representadas por uma Matriz de Atributo e Valor (AVM). Cada atributo pode ter, no máximo, um valor. Quando duas AVMs se unem, mapeiam uma nova AVM, a qual reúne os atributos e valores das duas AVMs originais.

Fried e Ostman afirmam que a tarefa dos princípios de unificação é assegurar que atributos com valores contraditórios falhem ao se combinar (2004). Os autores relembram a definição de Fillmore e Kay (1995 apud FRIED & OSTMAN, 2004) acerca de unificação: uma operação pela qual duas ou mais AVMs são combinadas de modo que a AVM resultante contém exatamente o que está presente em seus AVMs componentes.

O exemplo da Figura 15 mostra a combinação das palavras *much* e *snow*.

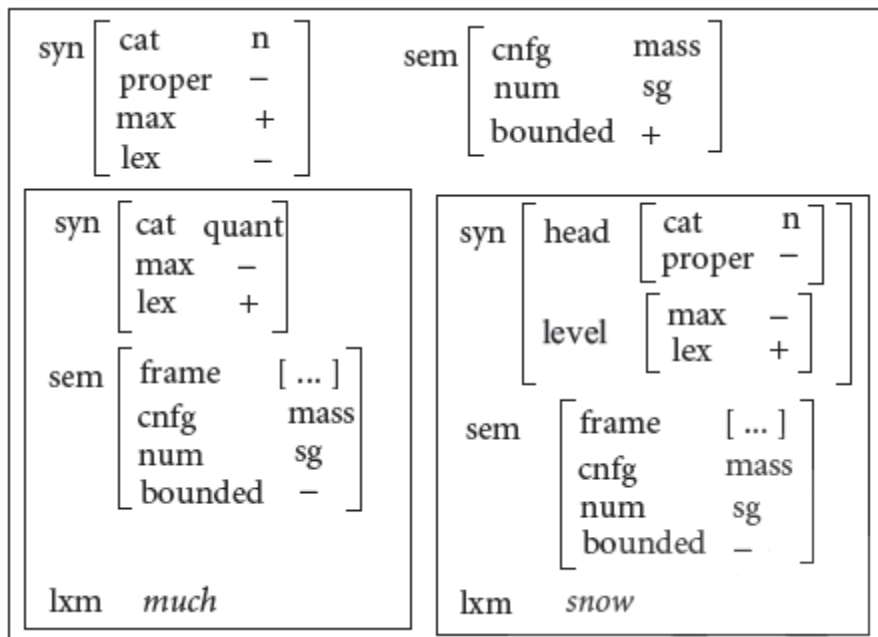


Figura 15 (FRIED & OSTMAN, 2004, p.34)

Os autores ressaltam que os atributos que precisam unificar nesses construtos podem ser descritos como traços semânticos que são projetados sobre o sintagma como um todo e constituem a base da concordância gramatical entre constituintes neste padrão específico. No exemplo acima, o construto *much snow* tem suas especificações semânticas combinando nos três atributos – cnfg (configuração), num (número) e bounded (delimitação) –, isto é, esses valores semânticos se unificam.

Por outro lado, se tentarmos combinar *much* e *book*, fracassaremos:

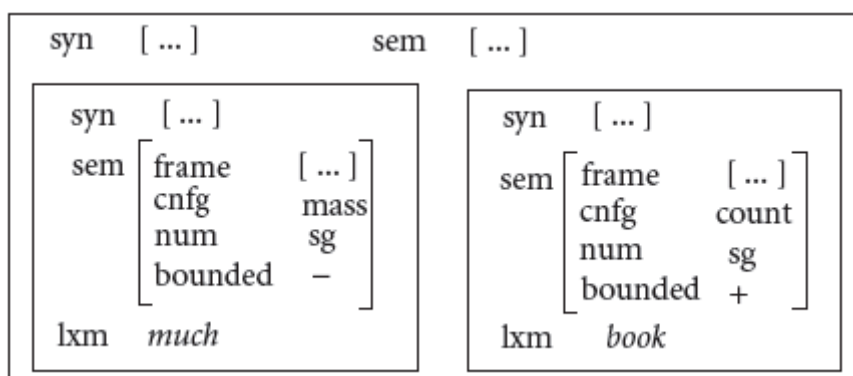


Figura 16 (FRIED & OSTMAN, 2004, p.35)

Isso ocorre em função de valores conflitantes em dois atributos semânticos. Como configuração e delimitação são marcados com valores diferentes, já que

much quantifica massas não delimitadas e *book* é um objeto delimitado, é impossível se obter uma unificação bem sucedida.

O formalismo presente nessa vertente da Gramática das Construções é mais especificado e explícito – o que justifica o fato de muitos de seus trabalhos serem voltados à área da Linguística Computacional. Por outro lado, segundo Goldberg, uma abordagem baseada em unificação falha ao captar propriedades semântico-lexicais detalhadas. A autora explica que, embora a Gramática das Construções baseada em Unificação use constantes para representar significados semânticos, o que evita a decomposição do significado em um conjunto fixo de traços, o formalismo presente supervaloriza elementos sintáticos, os quais são os mais recorrentes, e são os traços recorrentes os mais relevantes para o mecanismo da unificação (GOLDBERG, 2006, p.216).

2.3 O Constructicon

Para o que a FrameNet se propõe – a descrição da valência de unidades lexicais –, sua metodologia de análise a satisfaz. Contudo, as estruturas gramaticais utilizadas na descrição dessas unidades lexicais são limitadas para relações de predicação, modificação e complementação, como apontam Fillmore et al. (2012). As propriedades sintáticas e semânticas de algumas sentenças não podem ser totalmente capturadas pela abordagem dada pela FrameNet.

Nesse contexto, surge a necessidade de um novo recurso, desenvolvido para atender justamente aos casos periféricos da língua, abarcando estruturas mais complexas e idiomáticas, o Constructicon. Tal recurso configura-se em uma tentativa de se aplicar a interface e os princípios de anotação desenvolvidos no âmbito da FrameNet às construções, conforme são analisadas pela SBCG, ou por outras vertentes da Gramática das Construções baseada em Unificação. Assim sendo, as camadas usadas na anotação de sentenças representam a unificação de um conjunto de valores para os atributos Elementos da Construção, Tipo Sintagmático, Função Gramatical (etc.), associados ao próprio construto.

Fillmore, autor do projeto Construction, o define como um registro de construções gramaticais no qual sentenças são anotadas observando-se quais

de suas partes são licenciadas por construções, e quais são essas construções (FILLMORE, 2008a). O Constructicon tem por objetivo descrever as possibilidades sintáticas, semânticas e pragmáticas dessas construções gramaticais através de anotação computacional de sentenças, e exibir os resultados dessa anotação.

Conforme mencionado acima, o Construction surge para suprir a necessidade de análise de determinadas estruturas linguísticas não processáveis no modo lexicográfico de anotação. Na FrameNet, somente as valências das unidades lexicais são anotadas. Eventualmente, são também anotadas categorias léxico-gramaticais. Fillmore (2008a) aponta os tipos de multipalavras (*multiwords*) tratáveis pela FrameNet; são algumas delas as construções de suporte, as quais são tratadas como evocando um *frame* conectado ao substantivo em vez do verbo (como, por exemplo, em: “Joana deu um grito”); e nomes transparentes (como em: “saco de pão”; “copo de leite”).

Entretanto, de modo geral, outras informações acerca do contexto linguístico da UL alvo não são aludidas pela anotação lexicográfica. Observe as sentenças (14), (15) e (16), a seguir, nas quais notamos o contexto de *habilidade negativa* sinalizado pelas formas sublinhadas:

- (14) Ana é **nova** prá ler esse livro.
- (15) Pedro não **aguenta** mais ouvir essa música.
- (16) Vocês ainda **toleram** essa situação?

Nesses casos, por exemplo, não há um modo consistente de se apontarem as características referentes ao contexto de habilidade negativa (FILLMORE et al, 2012).

Quando se trata de anotação de texto corrido, é ainda mais fácil encontrar sentenças cujas estruturas não se prestam a um tratamento meramente lexicográfico. É o caso da sentença (17), retirada do jornal The Economist:

(17) *For all the disappointments, posterity will look more kindly on Tony Blair than Britons do today*²⁷.

Nessa sentença, encontramos ao menos três construções que carecem de uma descrição, a saber, construção concessiva (presente em: “*For all the disappointments*”), colocação metafórica em *phrasal verb* (“*look kindly on*”), e construção comparativa de duplo foco (“*[posterity] will look more kindly on Tony Blair than [Britons] do [today]*”) (FILLMORE, 2008a). Nenhuma das três construções mencionadas são passíveis de ser tratadas via FrameNet.

Uma vez que a anotação lexicográfica considera apenas uma UL alvo, e todas as construções supramencionadas possuem, no mínimo, três unidades alvo, já se torna improvável esse modo de anotação. Além disso, as estruturas alvo nem sempre são fixas, isto é, podem-se variar alguns aspectos nas construções citadas, como em “*For all the disappointments*”, em que “*disappointments*” pode ser trocado por outros substantivos, ou mesmo por orações. No modo lexicográfico, não é possível anotar uma UL cujo constituinte possa variar.

Outro tipo de uso linguístico cujos aspectos sintático-semânticos não podem ser inteiramente capturados pela anotação levada a cabo na FrameNet é a Construção “*What’s X doing Y?*” (WXDY), descrita por Kay e Fillmore (1999) e que pode ser observada na sentença (18):

(18) O que esse cachorro está fazendo em cima da minha cama?

Se tomássemos o verbo “fazer” como a unidade lexical alvo da sentença, poderíamos tentar anotá-lo como evocando o *frame* de *Agir_intencionalmente*, o que nos renderia uma análise semelhante à proposta em (19):

(19) ???[O que_{Ação}] [esse cachorro_{Agente}] está **fazendo** [em cima da minha cama_{Lugar}]

²⁷ Apesar dos desapontamentos, a posteridade vai olhar com mais simpatia para Tony Blair do que os britânicos o fazem hoje.

O fato de a análise supramencionada não abarcar o significado da sentença se explica pelo fato de que uma resposta como a que está presente na sentença (20) poderia enfurecer o enunciador.

(20) Acho que ele está dormindo.

A fúria, nesse caso, decorreria do fato de que, ao utilizar-se da construção WXDY, o falante não está, tão somente, propondo uma pergunta. Na realidade, sua intenção é enfatizar sua irritação com um fato – no caso, o de o cachorro estar deitado em sua cama – visto por ele como negativo e/ou inaceitável.

Esses dados semântico-pragmáticos não podem ser apreendidos por uma análise via padrão de valência, uma vez que eles não são evocados pela unidade lexical “fazer”, mas pela construção WXDY. Por essa razão, o Constructicon deve abranger conhecimento linguístico que excede a valência simples de palavras simples (FILLMORE, 2008a).

Embora ambos tenham por intuito descrever as construções de uma língua, o Constructicon tem objetivos diferentes daqueles que uma Gramática das Construções possui. Em alguns casos, Fillmore et al. (2012) comentam que serão oferecidas propostas precisas para o tratamento de uma construção, como apareceria em uma gramática; em outros, porém, as descrições serão apenas observações organizadas sobre construções individuais; em todos os casos, porém, pretende-se que o Constructicon contenha informações úteis para atividades em PLN.

De modo mais específico, o Constructicon visa a descrever construções em termos de suas propriedades gramaticais e seu potencial semântico. Além disso, seu objetivo é realizar tais descrições de modo compatível com os pressupostos teóricos da SBCG.

Fillmore et al. definem construções como sendo

(...) as regras que licenciam ‘novos’ signos linguísticos baseados em outros signos linguísticos. As estruturas licenciadas por uma ou mais construções são chamadas CONSTRUTOS, seguindo-se a terminologia da SBCG. Uma construção pode ser descrita formalmente, na forma de Matrix de Atributo e Valor, ou informalmente, em prosa (...) (2012, p.321) ²⁸.

²⁸ (...) the rules that license ‘new’ linguistic signs based on other linguistic signs. The structures licensed by one or more constructions are called CONSTRUCTS, following the terminology of Sign-based Construction Grammar. A construction can be described formally, in Attribute-Value Matrix form, or informally in prose (...)

Os autores ressaltam que a anotação precisa voltar-se aos construtos, captando as propriedades de um construto específico, em relação a uma construção específica que o licencia. Fillmore et al. (2012) explicam que o procedimento de anotação identifica e classifica os sintagmas, os quais indicam trechos da língua que instanciam as construções em questão. Na sentença seguinte, o trecho é sinalizado pelas chaves:

(21) *They hired me at {four dollars an hour}*

Indicam, também, os segmentos desses trechos que são determinados pelas exigências da construção, os quais estão sinalizados pelos colchetes na sentença (22):

(22) *They hired me at { [four dollars] [an hour] }*

E, ainda, apontam para elementos do contexto cujas propriedades exigem uma instância da construção dada, como a preposição fora das chaves em:

(23) *They hired me [at] { [four dollars] [an hour] }*

Seguindo padrão de anotação derivado da FrameNet, o Constructicon também apresenta uma anotação em camadas. Duas segmentações devem ser feitas para fins de anotação construcional: primeiramente é preciso isolar o trecho sob o qual incide a construção e, a seguir, identificar as entidades linguísticas presentes no trecho, as quais representam os elementos constituintes da construção. O trecho inteiro corresponde à Mãe, ao passo que os componentes são as Filhas, ou Elementos da Construção (EC) – seguindo-se uma nomenclatura paralela àquela usada na FrameNet para os Elementos de Frame. Assim, se considerarmos a construção da Velocidade, o sintagma “80 Km por hora” tem dois ECs:

[80 Km] [por hora]

Um Tipo Sintagmático (TS) é atribuído para cada EC, por exemplo, Sintagma Nominal (SN) e Sintagma Preposicional (SP). A camada do TS não implica na existência da camada da Função Gramatical (FG), como ocorre na FrameNet, sendo que esta última só será anotada no Constructicon caso seja de relevância para capturar generalidades da construção em análise.

[_{SN} 80 Km] [_{SP} por hora]

Os dois ECs juntos formam o construto da Velocidade de tipo SN. Essa informação será indicada na camada do TS do construto (TSCstr). A extensão da etiqueta indica a extensão do construto, e o valor da camada (“SN”) sinaliza a categoria sintagmática do construto:

[_{SN} [_{SN} 80 Km] [_{SP} por hora]]

Em algumas ocorrências, uma construção pode ser sinalizada por um material lexicalmente fixo especificado por essa construção. A camada do EEC tem como função demarcar tais palavras.

Através da Figura 17, Fillmore et al. (2012, p.11) ilustram como uma construção é anotada no Constructicon.

Layer	C	o	n	s	t	a	n	c	e	s	q	u	e	e	d	h	e	r	w	a	y	d	o	w	n	t	h	e	p	l	a	t	f	o	r	m
CE	T	h	e	m	e					T	r	a	n	s	i	t	i	W	a	y	_	e	x	p	P	a	t	h								
CE																			T	h	e															
GF	E	x	t																							D	e	p								
PT	N	P								V	P	f	i	n				N	P						P	P										
CEE																																				
CstrPT										V	P	f	i	n																						
GovX																																				

Way_means - CE	Way_means - CE	Way_means - GF	Way_means - PT	Way_means - CEE	Way_means
Direction <F8>		Path <F5>			
Goal <F7>		Source <F6>			
Intransitive_mean...		Theme <F1>			
Manner <F9>		Transitive_means...			
Means <F10>		Way_expression <...			
Modifier <M>					

Figura 17: Anotação de uma instância da Construção Way no Constructicon da FrameNet

Como os autores explicam, o construto compreende o trecho “*squeezed her way*”. Na primeira linha, “*squeezed*” é classificado como meio transitivo para indicar que o verbo indicando a forma como o tema percorreu o caminho geralmente tem um argumento que não é exprimível na construção. “*Her way*”, por sua vez, é identificado como a EXPRESSÃO_WAY. Não há preenchimento da camada FG (GF). O TS (PT) é assinalado como SV finito para “*squeezed*” e SN para “*her way*”. A Construção Way possui material lexicalmente especificado e, por isso, o trecho “*her way*” é marcado na camada EEC (CEE). A camada TSCstr (em inglês, CstrPT) revela que se trata de um sintagma verbal finito.

No restante da sentença, dois ECs (CE) são encontrados: Tema e Trajetória, marcados na camada FG como Ext e Dep, respectivamente; e, na camada TS, SN e SP, respectivamente. O Tema assinalado logo abaixo de “*her*” indica que o pronome possessivo identifica-se e deve concordar com o argumento externo (Fillmore et al., 2012, p.11).

Fillmore et al. (2012) descrevem os tipos de construções que devem constar em um Constructicon; são elas: construções que introduzem um *frame*, construções que aumentam a valência, construções sem significado, construções determinadas contextualmente, construções exocêntricas e sem núcleo, construções de “enchimento” e construções definidoras de tipos oracionais. Vejamos as características de cada uma delas a seguir.

As **construções que introduzem um *frame*** referem-se a certas construções que licenciam expressões evocadoras de *frames*. Essas expressões atuam de modo análogo a palavras que evocam *frames*. Como exemplo, em português, temos a Construção Modal com *Dar*, a qual será definida no capítulo 3 e que também já foi estudada por Salomão (1990; 2007b). Essa construção, que pode ser exemplificada por sentenças como “Dá pra dormir na viagem?”, evoca o *frame* de Possibilidade.

Construções que aumentam a valência são aquelas capazes de ampliar a valência de um signo. Como exemplo, têm-se as construções comparativas, capazes de aumentar a valência de adjetivos. Por exemplo, se, por um lado, o adjetivo “bonito” não prevê a ocorrência de qualquer argumento em sua valência em sentenças como “Pedro é bonito”; por outro, a Construção

Comparativa de Superioridade aumenta sua valência, exigindo a ocorrência de um argumento que represente o indivíduo em relação ao qual Pedro é mais bonito, como na sentença “Pedro é mais bonito que o irmão”.

Embora construções sejam definidas como um pareamento de uma forma a um significado, existem pesquisadores que questionam se, de fato, todas as construções evocam *frames*. Enquanto Goldberg defende a semântica inerente a todas as construções, Fillmore, por exemplo, postula que existem construções puramente formais – como a construção da inversão do sujeito e auxiliar, da língua inglesa. Assim sendo, o Constructicon assume a legitimidade das construções semanticamente nulas e, por isso, descreve as **construções sem significado**.

As **construções determinadas contextualmente** aludem aos casos em que determinadas construções criam expressões cujos requisitos contextuais não são determinados por seu significado. Os autores exemplificam esse tipo de construção com construtos como “meu primeiro amor”, a partir do qual depreende-se uma relação entre o possuidor do amor e a pessoa amada diferente da interpretada em “meu amor”. No primeiro caso, que é uma instância da Construção de Unicidade, tem-se que “meu primeiro amor” refere-se à situação em que o enunciador amou alguém pela primeira vez. Entretanto, em “meu amor”, não há necessariamente a interpretação de que a expressão se refere a uma situação em que o enunciador amou alguém, podendo se referir, exclusivamente, à pessoa amada (FILLMORE ET AL., 2012, p.15).

Uma construção pode licenciar estruturas com propriedades sintáticas não previsíveis a partir dos atributos de suas partes. Essas construções podem ser descritas como **exocêntricas** ou como construções que não têm um núcleo sintático definido. Exemplos desse tipo de construção são expressões como “mês passado” ou “esta semana”, as quais têm valor adverbial como um todo, mas são compostas por núcleos nominais (FILLMORE ET AL., 2012, p.15).

Nas **construções de “enchimento”**, apresentam-se construtos com um único signo filha, possivelmente uma única palavra, e um signo mãe com uma forma relacionada ou idêntica à filha, e um significado que incorpora o significado do signo filha. Tais construções são capazes, por exemplo, de alterar propriedades dos nomes, transformando-os de incontáveis para

contáveis, em casos como “me dá um dinheiro?”, em que o nome incontável “dinheiro” passa a ser precedido pelo artigo “um”.

Por fim, as **construções definidoras de orações** referem-se às construções mais genéricas e mais produtivas de uma língua, tais como a Construção Imperativa, a Construção Transitiva e assim por diante.

Definidos o propósito de um Constructicon, sua metodologia de anotação e os tipos de construção que deve conter, passamos, no capítulo seguinte a analisar em que medida as construções em *Para Infinitivo* definidas por Torrent (2009; 2011; no prelo) podem ser incluídas nessa base de dados, ou se podem ser tratadas como padrões de valência de ULs evocadoras de *frames* específicos.

3 AS CONSTRUÇÕES DA FAMÍLIA PARA INFINITIVO

No presente capítulo, apresentaremos as construções da família *Para Infinitivo*, as quais serviram de objetos de análise no processo de implantação do Constructicon para o português brasileiro. Nossa escolha pela família *Para Infinitivo* partiu da tese de Torrent (2009), na qual tais construções são descritas e analisadas sob a ótica da Gramática das Construções de Goldberg (1995, 2006). Na tese, foram descritas 22 construções em *Para Infinitivo*, mas essas análises foram refinadas e expostas em trabalhos mais recentes (TORRENT, 2011; no prelo)²⁹.

Como veremos no transcorrer deste capítulo, as construções da família *Para Infinitivo* apresentam uma miríade de possibilidades no que diz respeito à tipologia construcional. Por essa razão, optamos por implantar o Constructicon descrevendo as construções dessa rede. Há construções de estrutura argumental bem aberta – como a Construção de Dativo com Infinitivo –, outras que dependem de classes verbais ou adjetivais específicas – como a Construção Manipulativa –, e, ainda, construções lexicalmente especificadas – como a Construção Volitiva. Assim, a família *Para Infinitivo* se mostrou ideal para iniciarmos o Constructicon, visto que diferentes possibilidades poderiam ser testadas e observadas na anotação.

Torrent (no prelo) propõe uma rede de construções de disposição radial, na qual são descritos o protótipo da rede e os links entre seus membros. Os links de herança utilizados nas análises de Torrent são baseados na proposta de Goldberg de herança construcional de modo normal (ver subseção 2.2.1.2).

Esse conjunto de construções pode ser chamado de *família*, uma vez que seus elementos compartilham propriedades sintático-semânticas. Com relação à forma, as construções dessa família compartilham o esquema sintático [SN¹ V SA/SN² para (SN³) V_{INF}] ou são a ele relacionadas através de links de herança. Na construção, como o SN³ pode ou não constar, sua representação se dá entre parênteses – o verbo infinitivo sempre terá um sujeito, mas esse poderá ser explícito ou inferido. Todas as construções da

²⁹ Ao longo desta seção, será feita referência aos *frames* evocados pelas construções da família *Para Infinitivo*. O leitor poderá ter acesso à descrição de tais *frames*, bem como às construções, através da guia Dados do sítio da FrameNet Brasil: <http://www.framenetbr.ufjf.br>.

família herdam propriedades da Construção Adjuntiva de Finalidade, representada formalmente em [para (SN) V_{INF}].

No que diz respeito ao sentido, Torrent define do seguinte modo a propriedade partilhada por essa família:

(...) todas as construções em *para (SN) infinitivo* apontam para um espaço mental que representa, ancorado no esquema cognitivo do movimento em direção a um alvo, uma resultante virtual desejada de uma dada ação, expressa por uma construção de estrutura argumental, ou pelo contexto (TORRENT, 2009, p.75).

Ao dizer que as construções desse esquema sintático revelam em seu significado “uma resultante virtual desejada de uma dada ação”, a análise aponta para o *frame* evocado. Assim, dizemos que as construções em *Para Infinitivo* compartilham o *frame* de Finalidade, ou algum *frame* a ele relacionado.

Conforme aponta Torrent (no prelo), a gramática tradicional assinala que a preposição “para” codifica o destino de um movimento ou o recipiente de uma entidade transferida. A relação da família *Para Infinitivo* com o *frame* de Finalidade também está relacionada a essa preposição, a qual é tradicionalmente classificada como aquela que introduz orações infinitivas de finalidade.

Torrent (no prelo) explica que as construções aqui descritas estão organizadas em uma rede de motivações e dispostas em uma hierarquia de lexicalidade-esquematicidade. A construção mais esquemática é definida em relação ao sentido geral da rede alvo pareado à sintaxe mais esquemática. Já as construções de classe específica são definidas quando há um sentido mais específico pareado a uma classe específica de verbo ou adjetivo que preencha, respectivamente, o verbo finito ou os sintagmas adjetivos. Construções especificadas lexicalmente, por sua vez, ocorrem apenas quando verbos ou adjetivos específicos preenchem o verbo finito ou os sintagmas adjetivos, apresentando significado mais fortemente não composicional.

A hierarquia anteriormente descrita pode ser observada na Figura 18 (TORRENT, no prelo). O diagrama representa as construções dispostas conforme o grau de lexicalidade-esquematicidade mencionado acima, de modo que, as construções mais esquemáticas encontram-se no topo do diagrama, enquanto as mais lexicalizadas estão em sua base. As setas indicam os links

de herança relacionando as construções. As letras S, M, e I sinalizam o tipo de link de herança presente, a saber, herança por Subparte, por Extensão Metafórica e por Instanciação, respectivamente.

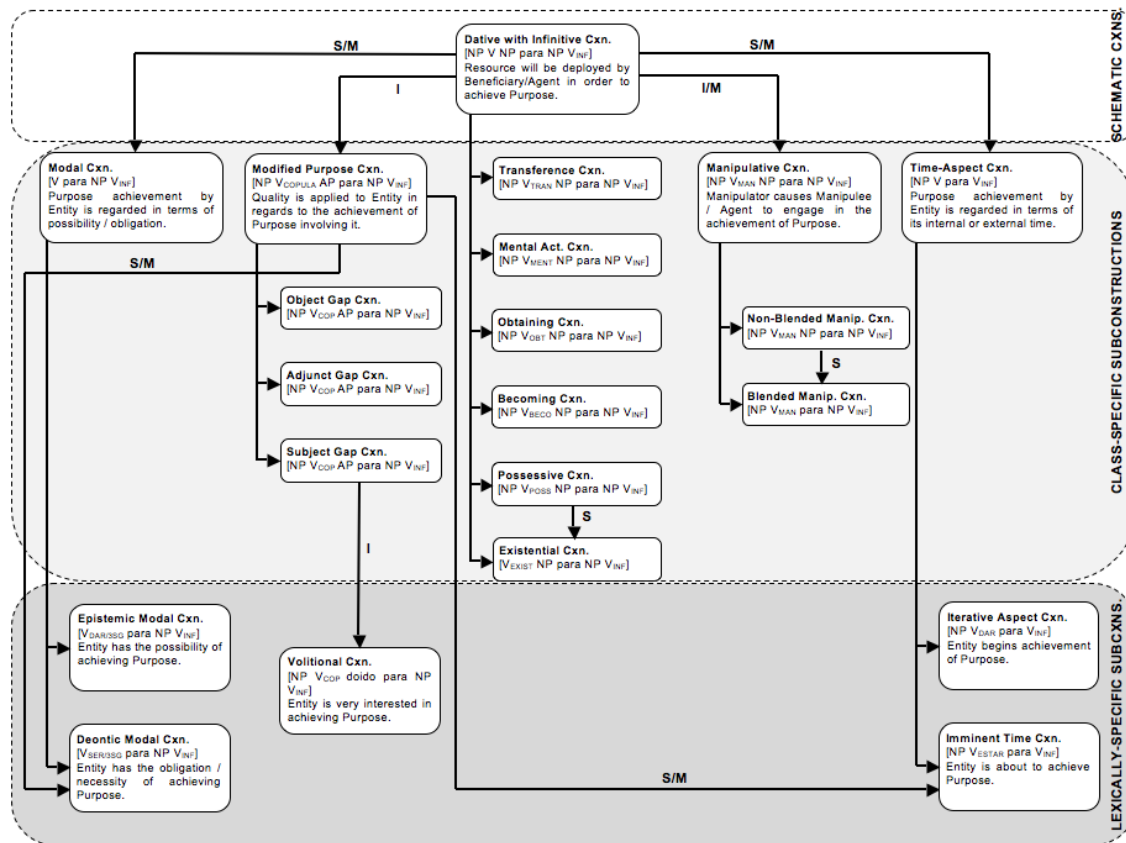


Figura 18: Hierarquia de lexicalidade-esquematicidade da Família *Para Infinitivo*

A **Construção de Dativo com Infinitivo** configura-se como a mais esquemática da família *Para Infinitivo* – o que pode ser observado na Figura 18 – ou seja, trata-se da construção prototípica da rede. Seguindo-se a análise de Salomão (1990, 2007b) acerca da construção de Habilitação com o verbo “dar”, a Construção de Dativo com Infinitivo herda a estrutura da construção de Transferência de Posse por um link de Extensão Metafórica, e se combina com a Construção Adjuntiva de Finalidade para motivar seu *frame* (TORRENT, no prelo).

Assim, o esquema “para infinitivo” preenche o espaço do argumento oblíquo da construção de Transferência. Essa disposição acaba por concentrar os papéis de Beneficiário (da transferência) e Agente no SN dativo.

A Construção de Dativo com Infinitivo evoca, portanto, o *frame* de Habilitação, em que um Recurso habilita um Beneficiário/Agente a atingir uma Finalidade. O modo como tal Recurso se torna disponível ao Beneficiário/Agente é determinado pela semântica do verbo finito. São muitos os verbos que aparecem nessa posição na construção de Dativo. Torrent (no prelo) os divide em grupos segundo sua classe verbal, conforme exemplificado no Quadro 4.

Classe Verbal	Verbo	Exemplo
Transferência	Dar	Juliana deu mil reais pra mim fazer o serviço.
	Ceder	Maucha cedeu a casa para a gente morar.
	Devolver	Marcela devolveu o livro pra nós lermos.
	Deixar	Minha mãe deixou o dinheiro pra você comprar pão.
	Colocar	Eu coloquei a pilha pro brinquedo funcionar.
	Passar	Bruno passou o número pra eu ligar.
Atividade Mental	conhecer	Adrieli conhece algum truque para tirar manchas?
	Saber	A senhora sabe alguma simpatia pra parar de fumar?
	Ver	Natália viu uns hotéis pra gente ficar no feriado.
	Ler	Eli lê muitas histórias pro filho dele dormir.
Obtenção	Comprar	Carina comprou uma entrada pra mim ver o show.
	Pagar	Ele pagou a dívida para o seu pai vender a casa.
	Arrumar	Ele que arrumou a condução pra nós ir.
Tornar-se	criar coragem	A senhora tem que criar corage pá senhora cantar pas alma.
	acontecer	O que que aconteceu pra ela ficar assim?
Posse	Ter	Eles têm dinheiro para comprar a casa?
	estar com	Daniela está com o ingresso pra ir ao teatro.
Existência	Ter	Tem alguma coisa para ver aqui?
	Existir	Existem recursos para financiar o projeto?
	Faltar	Faltam dois reais pra mim comprar a passagem.

	Aparecer	Apareceram dois meninos pra eu dar aulas.
	Rolar	Pai, rola um trocado pra eu sair?

Quadro 4: Verbos finitos que aparecem na Construção de Dativo com Infinitivo

Torrent (2011, no prelo) e Silva e Torrent (2012) comparam a Construção de Dativo com Infinitivo à Construção *Wherewithal* da língua inglesa. Na análise de Fillmore (2008b), o SN é considerado um recurso suficiente para se desempenhar a atividade indicada pelo complemento infinitivo. Deste modo, o sujeito do SV infinitivo é controlado e, assim, bloqueia a interpretação de que o SV seria uma oração relativa infinitiva cujo núcleo é o substantivo que representa o recurso. Observe-se o exemplo (24):

(24) Então você tem tempo pra mudar de roupa. (NURC-RJ)

O SN “tempo” é o recurso suficiente para que o Beneficiário/Agente “você” realize a finalidade indicada pelo complemento infinitivo “mudar de roupa”.

A partir da Construção de Dativo com Infinitivo, concebida como a construção prototípica da rede *Para Infinitivo*, outras construções a ela se conectam formando uma categoria radial e gerando efeitos de prototipia (TORRENT, no prelo).

A **Construção Manipulativa** é vinculada à de Dativo com Infinitivo através de um link de Extensão Metáforica. A metáfora que subjaz a essa relação é a Metáfora do Conduto. De configuração sintática extremamente similar, essas construções se diferenciam, nesse aspecto, apenas pela posição do verbo finito, que, na construção Manipulativa, deve ser ocupada por um verbo manipulativo – como, por exemplo, “*pedir*”.

Contudo, Torrent (2009; 2011; no prelo) mostra que os mapeamentos funcionais das duas construções são diferentes. Na Construção Manipulativa, o manipulador preenche o espaço do primeiro SN, o qual guia a entidade aludida no segundo SN a alcançar algum objetivo desejado. Geralmente, guia também o terceiro SN, já que os dois SNs costumam ser coreferenciais. Dessa forma, nas ocorrências da Construção Manipulativa, percebe-se a evocação do *frame* de Pedir.

A Manipulativa se subdivide em duas construções: *mesclada* e *não-mesclada*. A primeira herda da segunda por meio de um link de Subparte, visto que o manipulado referido pelo verbo finito se mescla ao agente do infinitivo no SN introduzido por “para” (TORRENT, 2009; 2011; no prelo). Veja os exemplos (25) e (26):

- (25) Ele pediu o filho pra ele ficar.
 (26) Às vezes ela pede para mãe escolher uma roupa para ela.
 (NURC-RJ)

A primeira sentença é um exemplo da Construção Manipulativa Não-Mesclada, em que temos o Manipulador no primeiro SN “Ele” guiando o Manipulado “o filho” a desempenhar a atividade aludida pelo SV_{inf} “pra ele ficar”. Na segunda sentença temos um exemplo da Construção Manipulativa Mesclada, herdeira da primeira, em que o manipulado se mescla ao agente do infinitivo em “para mãe”.

A **Construção de Propósito Qualificado** liga-se à construção de Dativo com Infinitivo por um link de Instanciação. Duas restrições podem ser encontradas na primeira construção: o verbo finito precisa ser preenchido por uma cópula, e um SA que modifica o primeiro SN deve ocupar o espaço após a cópula. Tais restrições justificam o link por Instanciação entre as construções.

A Construção de Propósito Qualificado se subdivide em três subconstruções cuja definição se estabelece em termos de qual argumento do verbo infinitivo deixa uma lacuna ao preencher a posição do primeiro SN. Se essa lacuna for a posição de objeto do verbo infinitivo, temos a Construção de *Gap* de Objeto, em que o adjetivo provavelmente caracterizará a finalidade com relação a seu grau de dificuldade. Por outro lado, se essa lacuna for a posição de sujeito, os adjetivos geralmente atribuem ao sujeito da cópula características relacionadas à sua habilidade, capacidade ou prontidão para alcançar a finalidade, configurando a Construção de *Gap* de Sujeito. Por fim, a Construção de *Gap* de Adjunto refere-se aos casos em que a lacuna é um adjunto de instrumento. Nesses casos, os adjetivos referem-se à utilidade desse instrumento para a realização da finalidade (TORRENT, no prelo). Observe-se os exemplos a seguir:

- (27) Esse exercício é fácil para você resolver.
- (28) Ela não estava pronta para viajar sozinha.
- (29) Esse sabão é bom pra tirar mancha.

No exemplo (27), temos a Construção de *Gap* de Objeto, já que o argumento que deixa uma lacuna é o objeto do verbo “resolver”. Esse objeto se torna o sujeito da cópula. Note-se, ainda, que a finalidade “para você resolver” é caracterizada quanto a seu grau de dificuldade – “fácil”. A Construção de *Gap* de Sujeito aparece no exemplo (28), visto que a lacuna deixada pelo sujeito do verbo infinitivo “viajar” é ocupada pelo sujeito da sentença “Ela”. O adjetivo “pronta” atribui característica ao sujeito da cópula “Ela” com relação à prontidão em se atingir a finalidade “pra viajar sozinha”. A sentença (29) apresenta uma instanciação da Construção de *Gap* de Adjunto. A lacuna é feita pelo adjunto relativo ao verbo infinitivo “tirar”. O adjetivo que ocupa o SA “bom” alude à utilidade do SN que o precede em executar a finalidade em “pra tirar mancha”.

Conectada à subconstrução de *Gap* de Sujeito por um link de Instanciação, a **Construção Volitiva** é um caso lexicalmente especificado da construção mãe, uma vez que o espaço do SA precisa ser ocupado por “doido” ou “louco”. Com esse novo traço, a construção ganha um sentido particular, qual seja, o de volição extrema – *frame* de Desejar. Como essa construção é lexicalmente especificada, aparece na base do diagrama representado na Figura 18.

- (30) Eu ficava louco pra voltar. (NURC-RJ)

Na sentença acima, vemos o SA sendo preenchido por “louco” e a finalidade de volição extrema representada pelo SV_{inf} “pra voltar”.

As próximas construções a serem apresentadas configuram-se como de domínio específico de classe (conforme diagrama da Figura 18) uma vez que elas também herdam propriedades e estrutura de construções mais esquemáticas que a Construção de Dativo com Infinitivo. Como Torrent esclarece (no prelo), a Construção Modal que iremos tratar nesse momento relaciona-se, também, à Construção Modal mais geral do PB.

A **Construção Modal** conecta-se à construção de Dativo com Infinitivo via link de Extensão Metafórica e, também, via link de Subparte. O verbo finito

pode ser preenchido pelo verbo “dar”, gerando um significado de modalidade epistêmica e evocando o *frame* de Possibilidade, ou pelo verbo “ser”, o qual origina o significado de modalidade deôntica e evoca o *frame* de Ser_obrigado. Observem-se os exemplos (31) e (32):

- (31) Então dá pra ir todo mundo. (NURC-RJ)
 (32) Então eu acho que frio é pra ficar dentro de casa, chuva também, (NURC-RJ)

Na primeira sentença, vemos o verbo “dar” preenchendo a posição do verbo finito. A frase poderia ser parafraseada do seguinte modo: “Então todo mundo pode ir”, o que evidencia o *frame* de Possibilidade sendo evocado. A segunda sentença, por sua vez, tem o verbo “ser” ocupando a posição de verbo finito, evocando junto ao VP_{inf} o *frame* de Ser_obrigado.

Finalmente, a **Construção de Tempo/Aspecto** herda da construção de Dativo com Infinitivo através de um link de Subparte e de Extensão Metafórica, o qual se sustenta na metáfora ESTADOS SÃO PONTOS EM UMA TRAJETÓRIA. Nessa construção há, também, duas subconstruções lexicalmente especificadas, a saber: a **Construção de Aspecto Iterativo**, a qual aponta que a ação indicada pelo verbo no infinitivo tem acontecido repetidamente; e a **Construção de Tempo Iminente**, que indica que o sujeito está prestes a alcançar a finalidade codificada pela oração infinitiva (TORRENT, no prelo).

- (33) O presidente deu para driblar seus seguidores de sempre como aconteceu quando embarcou para Juiz de Fora. (Cetenfolha)
 (34) A faxineira tá pra chegar.

Na sentença (33), temos a Construção de Aspecto Iterativo, em que a ação sinalizada pelo VP_{inf} “driblar” tem ocorrido de modo reiterado. Nessa construção o *frame* evocado é o de Iniciar_atividade_iterativa. A sentença (34) revela a Construção de Tempo Iminente. Nela vemos o sujeito “a faxineira” prestes a atingir a finalidade presente no VP_{inf} “chegar”.

Além de apresentar as construções que iniciarão o Constructicon brasileiro, este capítulo objetiva discutir qual tratamento cabe a cada uma

dessas construções considerando-se os objetivos de tal recurso sintático. Isto significa dizer que, apesar de Torrent (2009; 2011; no prelo) fornecer análises que sustentam o tratamento construcional dos vários padrões aqui discutidos no âmbito de uma Gramática de Construções, isso não significa que todas elas devam figurar no Constructicon, sendo necessário verificar quais construções podem ser tratadas via padrão de valência, e quais, por outro lado, devem ser incluídas no repertório de construções. É o que discutiremos nas seções seguintes.

É importante ressaltar que tratar uma determinada construção via anotação lexicográfica não significa afirmar que essa estrutura não seja uma construção. A decisão de incluir ou não uma construção no Constructicon é tomada tendo-se em vista o foco que ambos os projetos – FrameNet e Constructicon – têm na Linguística Computacional. Como os trabalhos são desenvolvidos principalmente para tarefas em PLN, isso faz com que as construções possam ser resolvidas de maneiras diferentes. Para tomar essa decisão, é preciso considerar a não-redundância de anotações no banco de dados, assunto que discutiremos no próximo capítulo.

3.1 Construções Tratáveis como Padrões de Valência

Dentre as construções apresentadas no início deste capítulo, apenas a Construção Manipulativa receberá um tratamento via padrão de valência. Vejamos a sentença (35).

- (35) Ela... era analfabeta... então pedia para a gente escrever o jogo...
(NURC-RJ)

A Construção Manipulativa evoca o *frame* de Pedir. Assim sendo, façamos uma análise em camadas, aos moldes da FrameNet, para a sentença (35) tomando-se como UL alvo o verbo “pedir” (Figura 19).

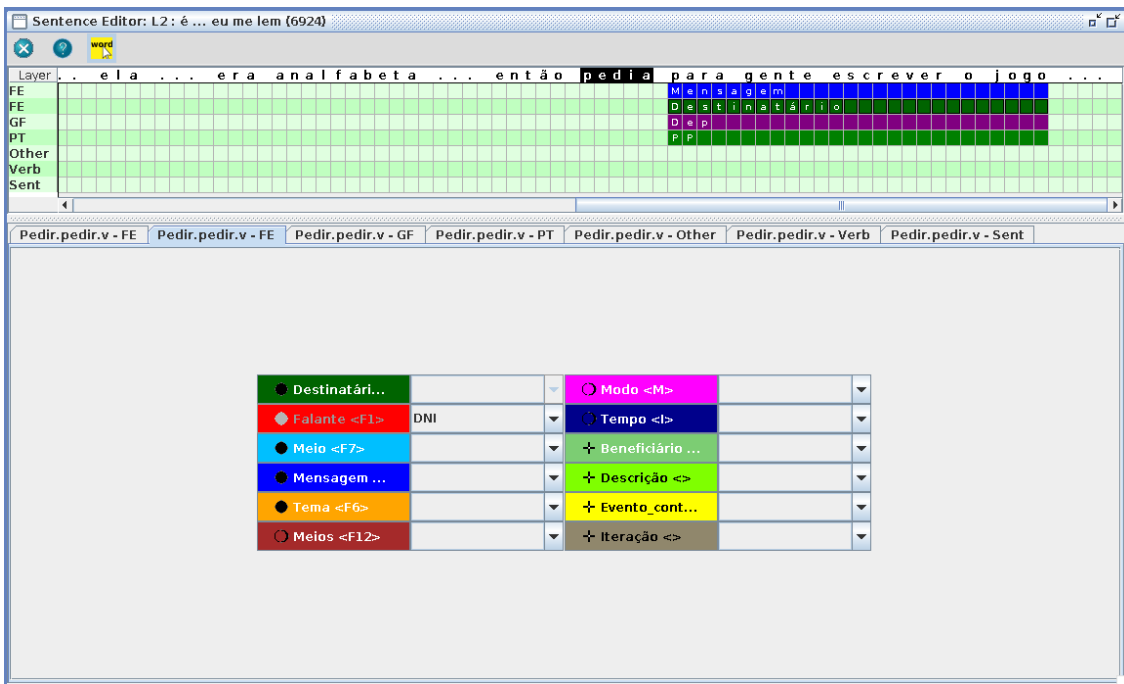


Figura 19: Anotação lexicográfica da UL “pedir”

Agora, observe a sentença (36) e sua análise em camadas:

(36) Por que a senhora não pediu uma carona? (NURC-RJ)

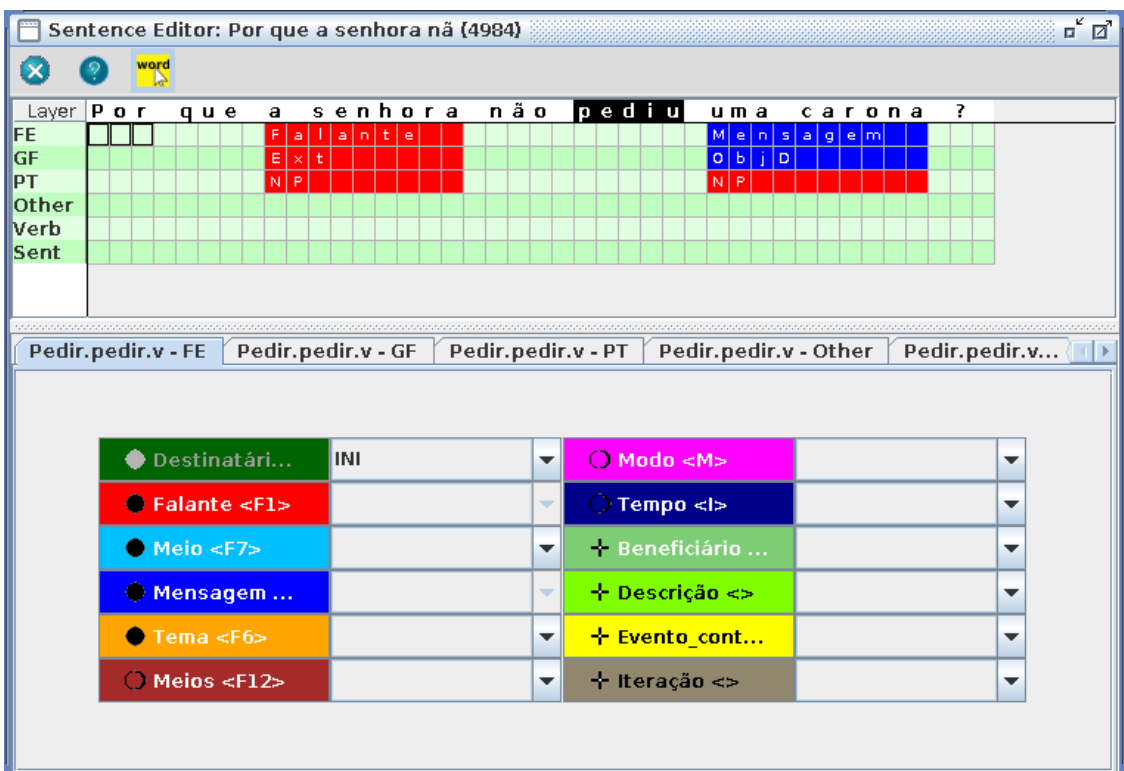


Figura 20: Anotação lexicográfica da UL “pedir”

Embora a estrutura argumental das sentenças varie – na primeira, o EF Mensagem, conflacionado³⁰ ao EF Destinatário, se realize como um SP e tenha função gramatical de Dep, enquanto na segunda o mesmo EF apareça como um SN e um DObj –, ambas evocam o mesmo *frame*, o de Pedir. Isso nos leva a concluir que tal *frame* é evocado pelo próprio verbo “pedir”, e não necessariamente pela construção.

Quando uma alteração na estrutura argumental gera uma grande alteração nos *frames* evocados por um mesmo verbo, isso sugere um tratamento construcional, uma vez que a evocação dos *frames* distintos não depende exclusivamente dos verbos, mas do padrão construcional. Na próxima seção trataremos desses casos.

Ademais, quando observamos os demais verbos atestados por Torrent (no prelo) na Construção Manipulativa, vemos que eles também evocam o *frame* de Pedir em mais de um padrão de valência, como pode ser visto nas sentenças de (37) a (42).

- (37) [Você_{Falante}] **convida** [os amigos_{Destinatário}] [pra tomar um drinque_{Mensagem}]. (NURC-RJ)
- (38) [Ele_{Falante}] fez a besteira de **convidar** [a turma toda_{Destinatário}] [para o casamento dele_{Mensagem}]. (NURC-RJ)
- (39) [Me_{Destinatário}] **avisa** [para eu poder, pra sair_{Mensagem}]. (NURC-RJ)
- (40) [A mulher do hotel_{Falante}] que **avisou** [que ia chover às duas horas..._{Mensagem}]. (NURC-RJ)
- (41) [Eu_{Falante}] ia **dizer** [pra você_{Destinatário} arrumar a sua cama_{Mensagem}]. (NURC-RJ)
- (42) Aí [eu_{Falante}] poderia **dizer** [pra você_{Destinatário}] [que eu acho lindos os campos_{Mensagem}]. (NURC-RJ)

Nas duas primeiras sentenças, a UL alvo é o verbo “convidar” e, embora na primeira, o EF Mensagem se configure como um [para SV_{inf}], e na segunda o mesmo EF seja apenas um SP, os *frames* evocados nas duas sentenças são o mesmo – o *frame* de Pedir. Nas sentenças (39) e (40), novamente, vemos uma variação no EF Mensagem, o qual se apresenta como um [para SV_{inf}] na sentença (39), e como uma oração completiva introduzida por “que” na sentença (40). Em ambos os casos, porém, vemos o mesmo *frame* de Pedir sendo evocado. As duas últimas sentenças apresentam o verbo “dizer” como

³⁰ EFs conflacionados são aqueles que se manifestam conjuntamente em um único sintagma.

UL alvo. O EF Mensagem encontra-se em um SV_{inf} na (41), enquanto na (42) trata-se de uma oração relativa. Novamente, não há mudança no *frame* evocado nas duas sentenças.

Isso se explica pelo fato de que, assim como se verificou com o verbo “pedir”, também os outros verbos da Construção Manipulativa são os responsáveis pela evocação do *frame*, e não a construção. Dessa forma, não há necessidade de anotar tal construção no Constructicon, uma vez que a anotação lexicográfica é suficiente para descrever os *frames* em questão.

3.2 Construções Incluídas no Constructicon

Nesta seção, além de apresentar as construções que não são analisáveis por padrões de valência, veremos porque razão isso ocorre. Começemos pela construção de Dativo com Infinitivo. Observe a seguinte sentença:

- (43) Então eles me davam aquele dinheiro para eu ir no bonde normal (NURC-RJ).

Analisando-a pelo viés lexicográfico, adotamos o verbo “dar”, presente na sentença, como a UL alvo. Assim, o *frame* evocado seria o de Dar, o que nos retornaria a anotação em camadas da Figura 21. Através dessa anotação, é possível localizar o que foi dado, quem recebe o dinheiro e por quem ele é dado, e com que finalidade o Doador dá o Tema ao Recipiente. Contudo, não se consegue depreender de parte alguma dessa análise o fato de que a construção presente na sentença evoca o *frame* de Habilitação, segundo o qual um Recurso habilita um Beneficiário/Agente a alcançar uma Finalidade. Isso ocorre porque não é o verbo “dar” que evoca o *frame*, mas a construção.

Embora o *frame* de Dar ofereça o EF Finalidade_do_tema e até se possa classificar [para SV_{inf}] como tal EF, se não for feita a descrição da construção, a análise como um todo perde em generalização. Como a Construção de Dativo com Infinitivo está presente em outros verbos além de “dar” – como vimos no início do capítulo –, e os *frames* evocados por esses outros verbos não apresentam EF similar, seria preciso anotar em cada *frame*

evocado pelas diferentes ULs o referido EF a fim de se captar tal nuance. Anotando a construção, ganha-se em generalização e economia – dois itens prezados em descrições linguísticas.

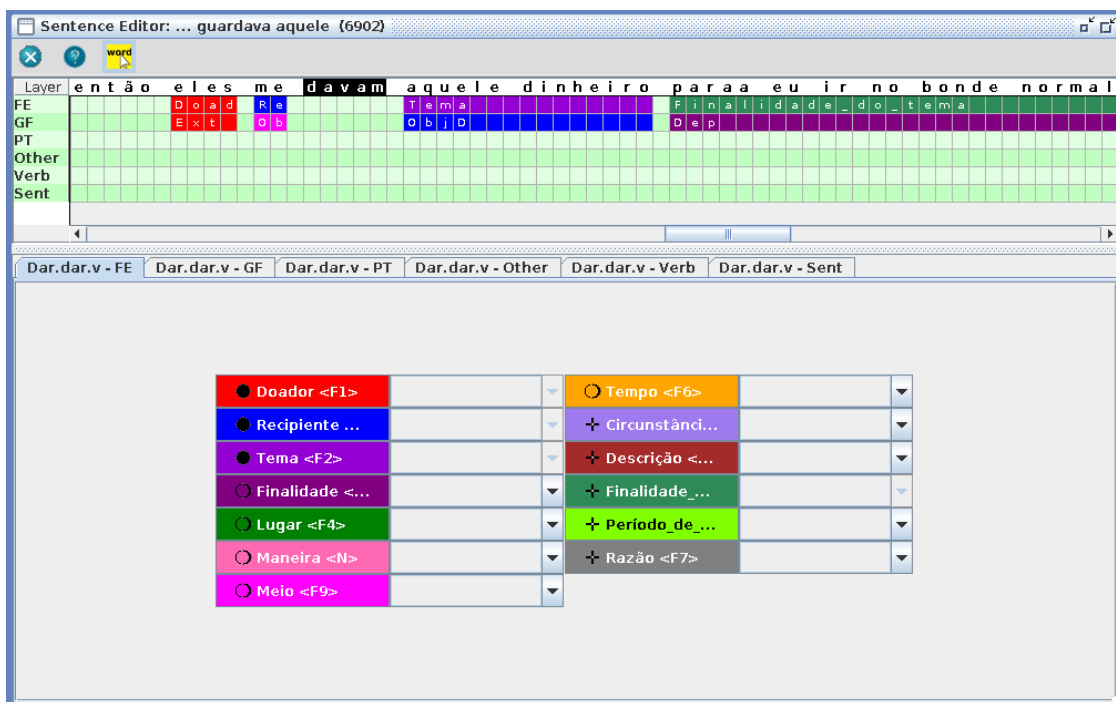


Figura 21: Anotação lexicográfica da UL “dar”

Se pensarmos em outros verbos que figuram na Construção de Dativo com Infinitivo, a necessidade de um tratamento construcional fica ainda mais evidente. Observe a Figura 22:

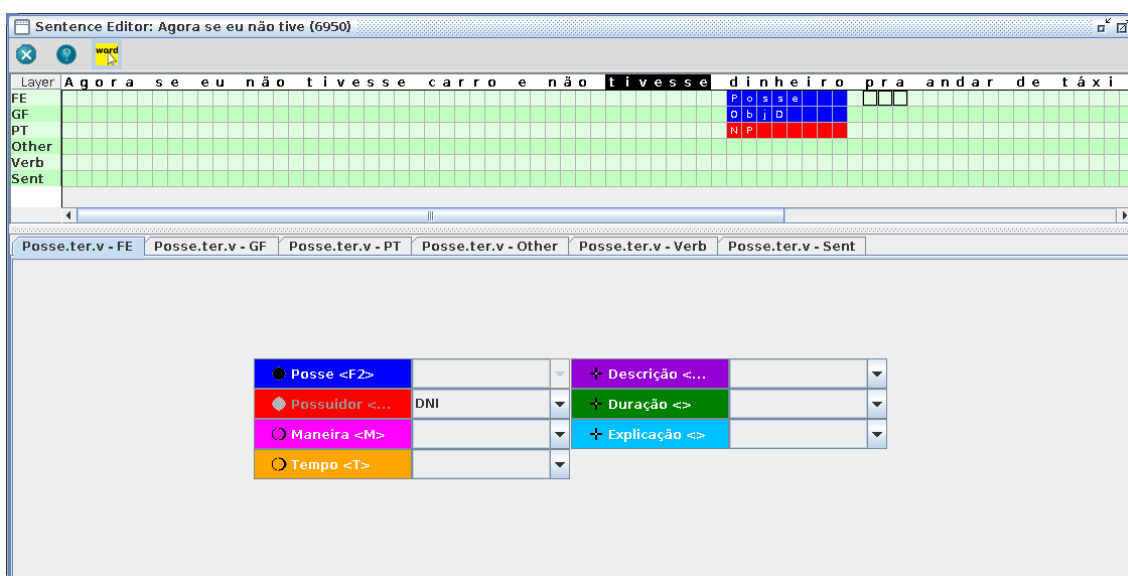


Figura 22: Anotação lexicográfica da UL “ter”

O verbo “ter”, tratado como a UL alvo da sentença, evoca o *frame* de Posse. Na análise em camadas, logo nos deparamos com o primeiro sinal de que o *frame* usado para descrever a sentença não recobre seu sentido de modo integral: não há sequer um EF previsto para descrever o evento indicado pelo sintagma “pra andar de táxi”.

Outra questão é que o *frame* evocado pela UL alvo, o de Posse, está longe de representar o *frame* que a construção evoca – um Recurso que habilita um Beneficiário/Agente a alcançar uma Finalidade. Somente um tratamento construcional pode realizar uma descrição plena da construção e do *frame* que essa evoca.

Observe, ainda, a Figura 23:

The image shows a software interface for sentence analysis. At the top, a window titled 'Sentence Editor: Foi uma representação [7026]' displays a sentence with a grid of colored cells representing different layers of analysis. The layers are: Layer, FE, GF, PT, Other, Verb, and Sent. The sentence 'eu comprei entrada pra ir assistir a uma determinada noite' is shown. Below the grid, there is a list of frame elements (EFs) for the 'Comércio_comprar.comprar.v' frame. The EFs are: Comprador, Mercadorias, Dinheiro, Finalidade, Lugar, Maneira, Meio, Seller, Taxa, Tempo, Unidade, Finalidade, Período, Razão, and Recipiente. Each EF is represented by a colored circle and a text label.

Figura 23: Anotação lexicográfica da UL “comprar”

O verbo “comprar”, tomado como UL alvo, é anotado no *frame* de Comércio_comprar. É possível identificar, na anotação lexicográfica, o Comprador, Mercadorias e, inclusive, a Finalidade_das_Mercadorias. Em outros termos, identificam-se os EFs que compõem o referido *frame*, o qual é, de fato, evocado, mas apenas pelo verbo. A construção evoca, como nos outros casos, o *frame* de Habilitação, o qual, além de não ser capturado pela descrição lexicográfica, forneceria uma descrição mais genérica e econômica das sentenças, evitando-se a existência de dois EFs tão semelhantes quanto Finalidade_do_tema e Finalidade_das_mercadorias.

Assim, a anotação dessas mesmas sentenças no Constructicon traria informações diferentes acerca de suas propriedades sintático-semânticas. Antes de passarmos à demonstração dessa anotação, entretanto, cabe fazer algumas considerações acerca da forma como o Constructicon de Berkeley define as construções e seus elementos.

Se tomarmos a Construção Way_manner da língua inglesa, teremos a definição ilustrada pela Figura 24.

Way_manner

Evokes the Motion frame.
Inherits Way_neutral.

- A verb exceptionally takes *one's way* (the CEE) as a direct object, where *one's* is a possessive pronoun coindexed with the external argument of the verb. Together, they indicate that some entity moves while performing the action indicated by the manner verb. The manner verb is either transitive or intransitive, and thus labeled either **Transitive_manner_verb** or **Intransitive_manner_verb**. Following *one's way* is an obligatory frame element indicating some core aspect of motion (Source, Path, Goal, Direction).
- The semantics of this construction is identical (or at least very close to) that of the frame Motion: A **Theme** moves under its own power from a Source, in a Direction, along a Path, to a Goal, by a particular means. In many cases the path traversed by the Self_mover is also created by them as they go, in a particular manner (i.e., while performing some temporally coextensive action) (as in *he whistled his way through the plaza*).
- [_{the} **She**] [_{t_man} **whistled**] [_{cee} **her way**] [_{Path} **down the lane**] [_{goa} **to the silo**].
- References:
 - Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
 - Kuno, Susumu and Takami Ken-ichi. 2004. *Functional Constraints in Grammar: On the Unergative-Unaccusative Distinction*. Amsterdam: John Benjamins Publishing Company.

- **CEE(cee)**: The construction-evoking element is the noun phrase *one's way*, where *one's* is coindexed to the **Theme**.
ex.: She whistled [_{cee} **her way**] down the lane to the silo. TRANSLATIONS
- **Direction(dir)**: The direction that the **Theme** heads in during the motion.
ex.: She whistled her way down the lane [_{dir} **towards the silo**]. TRANSLATIONS
- **Goal(goa)**: **Goal** is used for any expression which tells where the **Theme** ends up as a result of the motion.
ex.: She whistled her way down the lane [_{goa} **to the silo**]. TRANSLATIONS
- **Intransitive_manner_verb(i_man)**: The **Intransitive_manner_verb** takes the CEE as its object, and indicates the action performed by the **Theme** while it moves.
ex.: She [_{i_man} **whistled**] her way down the lane to the silo. TRANSLATIONS
- **Manner(man)**: Any expression which describes a property of motion which is not directly related to the trajectory or rate of motion expresses the frame element Manner. Descriptions of steadiness, grace, means of motion, and other things count as Manner expressions.
ex.: She [_{man} **gracefully**] whistled her way down the lane to the silo. TRANSLATIONS
- **Means(mea)**: An action which enables the **Theme** to move.
ex.: She laughed her way home [_{mea} **by not thinking about all the horrible things that had happened**]. TRANSLATIONS
- **Modifier(mod)**: The way in the **CEE** may be modified by an adjective. It often depicts a state of the **Theme** related to or resulting from their motion, but in some cases it modifies the path taken, motion event itself, or (rarely) a trait of the **Theme** unrelated to the fact that it is in motion. The **Modifier** is always indicated on the second layer.
ex.: She whistled her [_{mod} **cheerful**] way down the lane to the silo. TRANSLATIONS
- **Path(pat)**: The space traversed by the **Theme** between the **Source** and **Goal**.
ex.: She whistled her way [_{pat} **down the lane**] to the silo. TRANSLATIONS
- **Source(sou)**: **Source** is used for any expression which implies a definite starting-point of motion. In prepositional phrases, the prepositional object expresses the starting point of motion. With particles, the starting point of motion is understood from context.
ex.: She whistled her way [_{sou} **from the schoolhouse**] down the lane. TRANSLATIONS
- **Theme(the)**: The **Theme** moves, often under its own power. Normally it is expressed as an external argument.
ex.: [_{the} **She**] whistled her way down the lane to the silo. TRANSLATIONS
ex.: She whistled [_{the} **her**] way down the lane to the silo. TRANSLATIONS
- **Transitive_manner_verb(t_man)**: The **Transitive_manner_verb** takes the CEE as its object, and indicates the action performed by the Self_mover as it moves. It normally takes a post-verbal (i.e., non-subject) argument, but in this construction that FE is unexpressed and specified to be interpreted indefinitely (i.e., INI interpretation). Note that here "transitive" is meant to cover not just NP-taking verbs but also verbs that take PP, AP, AVP, etc. complements.

Figura 24: Definição da Construção Way_manner segundo o Constructicon de Berkeley

Conforme se pode notar na definição da Figura 24, há ECs que são definidos em termos sintáticos, tais como **Intransitive_manner_verb**, **Modifier** e **Transitive_manner_verb**. Outros assemelham-se a EFs, tais como **Goal**, **Manner** e **Theme**, por exemplo. O problema dessa confusão reside no fato de que ela não deixa claras as fronteiras entre informações relativas a um *frame* e

informações relativas a uma construção. Na própria definição da Construção *Way_manner* há a indicação de que sua semântica é idêntica ao *frame Motion*. Nesse caso, cremos que seria mais adequado, portanto, que ECs como Goal, Manner e Theme fossem definidos como EFs de Motion mapeados a ECs mais formais.

Com tais apontamentos, nota-se que o Constructicon americano, ao definir os ECs e as próprias construções, o faz caracterizando tanto propriedades formais, quanto semânticas, e, ainda mescla, em uma mesma definição, as duas informações. Isso é gerado pela ausência de critério na definição da construção, o que pode provocar confusão, sobretudo no que tange à compreensão do *frame* evocado pelo verbo, e aquele evocado pela construção.

Na esfera da FN-Br, como se propõe garantir a não-redundância do banco de dados – assunto sobre o qual discutiremos mais cuidadosamente no próximo capítulo –, os ECs são definidos exclusivamente em termos de suas propriedades formais. Essa resolução condiz com a proposta da Gramática das Construções baseada em Unificação, visto que, no caso de a construção evocar um *frame*, o EC será, então, pareado a um EF – se a construção for meramente formal, isso não ocorrerá.

Assim sendo, observe a definição da Construção de Dativo com Infinitivo realizada pela FN-Br (Figura 25) e, em seguida, a definição do *frame* que essa construção evoca (Figura 25):

Dativo_com_infinitivo

Definição [Definition]

Um VP_para_inf indica um propósito cuja realização depende de um recurso codificado em um Núcleo nominal ao qual o VP_para_inf se liga. Dada a interpretação habilitativa dessa construção, o VP_para_inf não é interpretado como uma oração relativa infinitiva comum.

Elementos da Construção [Construction Elements]

Núcleo[Head]

Nome que codifica um recurso empregado para a realização do propósito indicado pelo VP_para_inf.

VP_para_inf[VP_para_inf]

Oração que indica o propósito cuja realização é habilitada pelo recurso codificado pelo Núcleo. A construção de Dativo com Infinitivo bloqueia a interpretação deste VP_para_inf como uma simples oração relativa.

Exemplos de anotação [Annotation Examples]

-L Bom , o senhor não tem **dinheiro pra investir**, mas deve ter pelo menos um patrimônio pra segurar , não é

Figura 25: Definição da Construção de Dativo com Infinitivo dada pelo Constructicon da FN-Br

frames e a construções não interagem na plataforma da FN-Br, importada da FN de Berkeley. Assim, não é possível anotar os EFs Recurso e Propósito, relativos ao *frame* de Habilitação, que estariam pareados, respectivamente, aos ECs Núcleo e VP_para_inf. Porém, realizar o empreendimento de conectá-los está na agenda da FN-Br.

A construção que herda de Dativo com Infinitivo via link de Instanciação, a Construção de Propósito Qualificado, também necessita de uma descrição via Constructicon. Nesse caso, porém, há ainda mais lacunas sendo deixadas se a construção for tratada lexicograficamente. Vejamos porque isso ocorre nos três casos dessa construção. Observem-se os exemplos já mencionados no início deste capítulo:

- (44) Esse exercício é fácil para você resolver.
- (45) Ela não estava pronta para viajar sozinha.
- (46) Esse sabão é bom pra tirar mancha.

Se anotássemos a primeira sentença considerando “fácil” como UL alvo evocando o *frame* de Dificuldade, teríamos a seguinte anotação:

- (47) [Esse exercício_{Atividade}] é_{cop} **fácil** [para você_{Experenciador}] [resolver_{Atividade}].

Por essa anotação, é possível observar que a UL “fácil” evoca, de fato, o *frame* de Dificuldade, que também pode estar presente na Construção de *Gap* de Objeto. Contudo, essa análise não capta a relação de qualificação da Finalidade em “para você resolver” pelo Adjetivo “fácil”.

Por sua vez, a sentença (45) poderia ser anotada no *frame* de Atividade_preparada, tendo “pronta” como UL alvo:

- (48) [Ela_{Agente}] não estava_{cop} **pronta** [para viajar sozinha_{Atividade}].

Igualmente, aqui se observa o adjetivo “pronta” evocando o *frame* no qual foi anotado, o qual também pode ser evocado pela Construção de *Gap* de Sujeito. Entretanto, nada nessa anotação faz menção ao fato de que o Adjetivo “pronta” está em uma relação de qualificação com a Finalidade em “para viajar sozinha”.

Por fim, vejamos a anotação lexicográfica da sentença (46) no *frame* de Utilidade:

(49) [Esse sabão_{Entidade}] é_{cop} **bom** [pra tirar mancha_{Finalidade}].

Embora conste aqui um EF Finalidade, o qual não era oferecido pelos outros *frames*, ainda assim a relação de qualificação do Adjetivo “bom” em relação à Finalidade em “para tirar mancha” não é explícita.

Os adjetivos “fácil”, “pronta” e “bom” evocam *frames* específicos, mas sua relação com a Finalidade é dada pela Construção de Propósito Qualificado. Dessa maneira, tratar apenas as ULs adjetivais não é suficiente para captar a relação de qualificação da Finalidade pelo adjetivo no que tange à dificuldade/facilidade de sua realização, nem no que tange à utilidade do instrumento ou à capacidade do sujeito em realizá-la.

Ademais, em nenhuma das três anotações a questão do *Gap* foi abordada, nem sua relação com o Adjetivo em questão. Por mais esse motivo, a descrição construcional se faz necessária. Assim, novamente, ganha-se em generalização e economia. Veja-se, na Figura 30, a definição da construção de Propósito Qualificado pela FN-Br.

Propósito_qualificado

Definição [Definition]

Um sintagma nominal é qualificado por um Qualificador quanto à realização da Finalidade instanciada pela oração infinitiva regida por para, o VP_para_inf. Essa construção se caracteriza como um esquema mais genérico ao qual se vinculam as construções Gap_de_adjunto, Gap_de_objeto e Gap_de_sujeito.

Elementos da Construção [Construction Elements]

VP_para_inf[VP_para_inf] Sintagma verbal infinitivo regido por para, o qual codifica a ação que se quer realizar.

Qualificador[Qualifier] Adjetivo que qualifica o sintagma nominal quanto à realização da Finalidade instanciada pela oração relativa infinitiva regida por para.

Figura 30: Definição da Construção de Propósito Qualificado dada pelo Constructicon da FN-Br

A Construção Volitiva, por sua vez, embora herde da Construção de Propósito Qualificado, apresenta propriedades peculiares que diferenciam sua análise. Veremos que, ainda que a Construção Volitiva aparente ser facilmente tratada via padrão de valência, isso não procede. Para iniciar nossa discussão, observe as sentenças (50), (51) e (52) anotadas sob a ótica lexicográfica.

Trataremos “louco” como a UL alvo e os *frames* usados na anotação estão sinalizados ao fim de cada sentença:

- (50) [Eu_{Protagonista}] não sou_{Cop} **louco** [de voltar lá_{Comportamento}].
Frame de Propriedade_mental
- (51) ???[Eu_{Protagonista}] ficava_{Cop} **louco** [pra voltar_{Comportamento}]. (NURC-RJ)
Frame de Propriedade_mental
- (52) [Eu_{Experienciador}] ficava_{Cop} **louco** [pra voltar_{Evento}] (NURC-RJ).
Frame de Desejar

Na sentença (50), o *frame* de Propriedade_Mental consegue recobrir o sentido geral da sentença, uma vez que a UL alvo “louco”, de fato, evoca o *frame* referido. Nela, o Protagonista afirma que a ação codificada pelo sintagma “de voltar lá” é absurda a ponto de que realizá-la requer um grau de desvairo. Essa, porém, não é uma das construções da família *Para Infinitivo*.

A sentença (51), entretanto, tem sua análise comprometida ao ser tratada como uma instância do *frame* de Propriedade_Mental. Por essa análise, pode-se supor que o Protagonista contrai uma conduta insana para desempenhar o Comportamento em “pra voltar”, isto é, pela análise lexicográfica, só se pode depreender que o estado mental do Protagonista remete à loucura, o que não ocorre de fato.

O que a Construção Volitiva denota é que o Protagonista deseja de forma extrema realizar a atividade aludida no VP_{inf}. Toda a mudança semântica gerada nas sentenças (51) e (52) partiu de uma alteração na estrutura argumental da sentença (50), qual seja, a mudança da preposição “de” para “para” no sintagma posposto ao adjetivo. Assim, o *frame* de Desejar se mostra mais adequado para se anotar a sentença, como em (52).

Contudo, embora o *frame* de Desejar seja evocado na sentença (52), afirmar que o adjetivo “louco” evoca tal *frame* é um exagero. O *frame* de Desejar só é evocado por “louco” no padrão de valência da Construção Volitiva. Em outra estrutura argumental, o *frame* evocado será o de Propriedade_Mental. Portanto, o *frame* de Desejar é evocado pela Construção Volitiva, e não por “louco” apenas. Na Figura 31, observe a definição feita pela FN-Br acerca da Construção Volitiva.

Essa construção conta, ainda, com a camada do EEC presente em “louco para”, constituintes lexicalmente fixos e responsáveis por evocar a construção em questão.

Volitiva

Definição [Definition]

Um Desiderativo lexicalmente especificado como os adjetivos "doido" ou "louco" indica que um determinado sujeito deseja realizar a ação denotada pelo VP_para_inf.

Elementos da Construção [Construction Elements]

Desiderativo[Desiderative] Adjetivo que enfoca a ação indicada pelo VP como desejável. Nesta construção, o desiderativo é lexicalmente especificado como "doido" ou "louco".

VP para_inf[VP para_inf] Sintagma verbal infinitivo regido por para, o qual codifica a ação que se quer realizar.

Exemplos de anotação [Annotation Examples]

-R O menino deve ser **louco pra comer açúcar**

louco pra

Depois pra ir à escola, **louco pra ir à escola**, uma vontade pra ir, de ir paraa escola tremenda

louco pra

Os seus sessenta anos (pigarro) estava **louca pra se ver livre da filha**

louca pra

Figura 31: Definição da Construção Volitiva dada pelo Constructicon da FN-Br

A construção modal, herdeira da Construção de Dativo com Infinitivo, também necessita de um tratamento construcional. Na sua modalidade Epistêmica, notaremos que o verbo “dar” se faz presente assim como na Construção de Dativo. No entanto, na primeira construção ele ganha um sentido bastante peculiar, o qual também não se presta a uma anotação via padrão de valência. Acompanhe, o que acontece com a Construção Modal Epistêmica através do seguinte exemplo:

(53) Não dá pra ver não (NURC-RJ)

Tratar o verbo “dar” como a UL alvo dessa sentença nos faria anotá-la segundo o *frame* de Dar. Entretanto, esse *frame* não capta nenhum aspecto do *frame* de Possibilidade, evocado pela Construção Modal Epistêmica. Ao mesmo tempo, o *frame* de Dar não é evocado pela Construção Modal Epistêmica – não há nada sendo dado a alguém na sentença (54). Ao contrário, a sentença poderia ser transcrita na seguinte paráfrase: “Não é possível ver, não”, o que evidencia o *frame* de Possibilidade.

Veja-se, na Figura 32, a definição da Construção Modal Epistêmica pela FN-Br. Novamente, destaca-se a camada de anotação com o EEC “*dar para*”, marcada no exemplo de anotação.

Modal_epistêmica_dar

Definição [Definition]

Um Modalizador_epistêmico, necessariamente o verbo "dar", modifica um VP_para_inf que codifica uma ação ou evento possível.

Elementos da Construção [Construction Elements]

Modalizador_epistêmico[Epistemic_modal] Expressão que adiciona ao Vinf a noção de modalização epistêmica de possibilidade. Nesta construção, o Modalizador_epistêmico é especificado lexicalmente como o verbo "dar".

VP_para_inf[VP_para_inf]

Sintagma verbal infinitivo regido por para que codifica a ação ou evento possível.

Exemplos de anotação [Annotation Examples]

-R E dá pra ver o peixe na água

dá pra

A quantidade de ossos encontrada deu para encher quatro sacolas de supermercado

deu para

Figura 32: Definição da Construção Modal Epistêmica dada pelo Constructicon da FN-Br

No caso da Construção Modal Deôntica, a necessidade de um tratamento construcional também é evidente:

- (54) Então eu acho que frio é pra ficar dentro de casa, chuva também (NURC-RJ)

Nesse caso é nítida a necessidade de se analisar a sentença via Constructicon. A primeira questão que logo surge é decidir qual dos lexemas deve ser escolhido como UL alvo. Qualquer que seja escolhido não será capaz de abranger o *frame* em questão, o de Ser_obrigatório, uma vez que tal *frame* é evocado pela Construção Modal Deôntica, e não por uma unidade isolada. Veja: “frio” evoca o *frame* de Temperatura_Ambiente, “é” evoca o *frame* de Artistas_e_papéis e “ficar” evoca o *frame* de Residência. Nenhum dos *frames* evocados por essas ULs coincide com o *frame* evocado pela Construção Modal Deôntica, o que confirma que somente a anotação construcional será capaz de abrangê-lo. Observe-se a definição dessa construção pelo Constructicon da FN-Br (Figura 33).

Modal_deôntica_ser

Definição [Definition]

Um Modalizador_deôntico, necessariamente o verbo "ser", modifica um VP_para_inf que codifica uma ação ou evento obrigatórios / necessários.

Elementos da Construção [Construction Elements]

Modalizador_deôntico[Deontic_modal] Expressão que adiciona ao Vinf a noção de modalização deôntica de obrigação ou necessidade. Nesta construção, o Modalizador_deôntico é especificado lexicalmente como o verbo "ser".

VP_para_inf[VP_para_inf] Sintagma verbal infinitivo regido por para que codifica a ação ou evento obrigatórios / necessários.

Exemplos de anotação [Annotation Examples]

Apenas digo que são meus , que **é pra ele não vender**
é pra

Figura 33: Definição da Construção Modal Deôntica dada pelo Constructicon da FN-Br
 Por fim, considerando a Construção de Tempo/Aspecto, observemos a Construção de Aspecto Iterativo, presente na sentença (55):

- (55) O presidente deu para driblar seus seguidores de sempre como aconteceu quando embarcou para Juiz de Fora. (Cetenfolha)

Se tentássemos realizar uma análise via padrão de valência, logo nos depararíamos com a ausência de um EF capaz de descrever o papel desempenhado pelo sintagma “para driblar”, seguindo-se o *frame* de Dar. Afora isso, outra questão surge: a construção presente nessa sentença não faz menção a nada que tenha sido dado por alguém à outra pessoa, isto é, ao *frame* de Dar. De fato, a sentença assinala que “o presidente” tem driblado seus seguidores nos últimos tempos repetidamente.

Essas condições, portanto, apontam para uma análise construcional, já que a disposição argumental da construção gera uma grande alteração em relação ao *frame* evocado apenas pela UL “dar”. Confira a definição final proposta pela FN-Br a essa construção na Figura 34.

Aspectual_iterativa_dar

Definição [Definition]

Um Auxiliar_aspectual, necessariamente o verbo "dar", modifica um VP_para_inf que codifica uma ação ou evento cuja ocorrência se dá de maneira reiterada.

Elementos da Construção [Construction Elements]

Auxiliar_aspectual[Aspectual_aux] Expressão que adiciona ao Vinf a noção de ocorrência iterativa. Nesta construção, o Auxiliar_aspectual é especificado lexicalmente como o verbo "dar".

VP_para_inf[VP_para_inf] Sintagma verbal infinitivo regido por para que codifica a ação ou evento cuja ocorrência se dá de modo iterativo.

Exemplos de anotação [Annotation Examples]

O presidente **deu para driblar seus seguidores de sempre** como aconteceu quando embarcou para Juiz de Fora
deu para

Figura 34: Definição da Construção Aspectual Iterativa dada pelo Constructicon da FN-Br

Observe-se, no mesmo sentido, a Construção de Tempo Iminente:

(56) Eu estou pra colocar uma refrigeração aqui. (NURC-RJ)

O verbo “estar” tomado como UL alvo evoca o *frame* de *Artistas_e_Papéis*, que em nada se assemelha ao *frame* evocado pela construção, o de *Atividade_preparada*. O fato de que o SN “Eu” está prestes a realizar a ação indicada em “colocar uma refrigeração aqui” é evocado pelo esquema sintático da Construção de Tempo Iminente, a qual pode ter sua definição conferida na Figura 35.

Tempo_iminente_estar

Definição [Definition]

Um Auxiliar_temporal, necessariamente o verbo "estar", modifica um VP_para_inf que codifica uma ação ou evento cuja ocorrência é iminente.

Elementos da Construção [Construction Elements]

Auxiliar_temporal[Time_aux] Expressão que adiciona ao Vinf a noção de ocorrência iminente. Nesta construção, o Auxiliar_temporal é especificado lexicalmente como o verbo "estar".

VP_para_inf[VP_para_inf] Sintagma verbal infinitivo regido por para que codifica a ação ou evento cuja ocorrência é iminente.

Exemplos de anotação [Annotation Examples]

Bom , batida eu gosto , eu até hoje , eu acho que **está pra nascer** uma que eu não goste
está pra

Tem datas fixas , etc. onde as pessoas fazem certas coisas , há uma que **está pra chegar** , não é
está pra

Figura 35: Definição da Construção de Tempo Iminente dada pelo Constructicon da FN-Br

Como foi possível observar ao longo deste capítulo, foram definidas as construções da família *Para Infinitivo* e, em seguida, foi discutido quais delas requerem um tratamento construcional, e quais podem ser tratadas no modo lexicográfico de anotação. Demonstrou-se, também, o critério de se criarem os ECs com base em propriedades mais formais, apontando-se a necessidade de mapeá-los a EFs dos *frames* evocados pelas construções. Assim sendo, no próximo capítulo se faz necessário um debate sobre os critérios gerais que definem o que precisa ser anotado no Constructicon, bem como sobre inconsistências no modelo de anotação do Constructicon de Berkeley que precisam ser resolvidas quando de sua adaptação para o português brasileiro.

4 IMPLICAÇÕES DO TRATAMENTO COMPUTACIONAL DAS CONSTRUÇÕES DA FAMÍLIA PARA INFINITIVO

4.1 O Princípio da Não-Redundância

Os projetos FrameNet e Constructicon foram desenvolvidos com a finalidade principal de fomentar pesquisas e tarefas em PLN. O primeiro deles, a FrameNet, surge, conforme vimos anteriormente, como um recurso lexical. Suas anotações concernem Unidades Lexicais e os *frames* por elas evocados. À medida que o trabalho foi evoluindo, começou-se a perceber que uma série de estruturas linguísticas não podia ser incorporada a essa anotação, uma vez que esse recurso não fora desenvolvido para atendê-las. Nesse momento, a necessidade de se criar um Constructicon se fez presente.

A criação do Constructicon ocorre porque ambos os projetos visam a uma descrição completa da língua e, se uma das ferramentas não bastava para cobrir todos os casos encontrados, outra precisaria ser desenvolvida. Pensando nos possíveis usos dessas ferramentas em tarefas de PLN, um tradutor automático, por exemplo, é um recurso que deve abranger a língua de forma completa, não apenas as estruturas formalizáveis em termos dos padrões de valência prototípicos dos itens lexicais. Por isso é tão importante que os dois trabalhos ocorram de modo paralelo.

Entretanto, como vimos no capítulo três, algumas construções podem ser anotadas no modo lexicográfico, visto que o *frame* evocado será o mesmo, independentemente de sua estrutura argumental. Assim, se uma determinada estrutura já foi anotada na FrameNet, não é necessário ser anotada outra vez no Constructicon, isto é, não é preciso que uma mesma informação conste mais de uma vez no banco de dados. Isso é o que chamamos de Princípio da Não-Redundância.

A aplicação do Princípio da Não-Redundância se justifica devido a uma série de problemas gerados a partir de um armazenamento redundante. Para compreender melhor do que se trata tal princípio, antes são necessárias algumas definições. Como o Princípio se aplica sobre um banco de dados, iniciaremos por essa definição: “Um banco de dados é uma coleção de dados relacionados. Os **dados** são fatos que podem ser gravados e que possuem um

significado implícito” (ELMASRI & NAVATHE, 2010, p.4) [grifo original]. Há que se considerar, ainda, as seguintes propriedades implícitas a um banco de dados:

- Um banco de dados representa alguns aspectos do mundo real, sendo chamado, às vezes, de **minimundo** ou de **universo de discurso (UoD)**. As mudanças no minimundo são refletidas em um banco de dados.
- Um banco de dados é uma coleção lógica e coerente de dados com algum significado inerente. Uma organização de dados ao acaso (randômica) não pode ser corretamente interpretada como um banco de dados.
- Um banco de dados é projetado, construído e povoado por dados, atendendo a uma proposta específica. Possui um grupo de usuários definido e algumas aplicações preconcebidas, de acordo com o interesse desse grupo de usuários (ELMASRI & NAVATHE, 2010, p.4) [grifo original].

A geração de um banco de dados pode ser feita manualmente ou pode ser computadorizada. Para o segundo caso, é necessária, conforme Elmasri e Navathe (2010) assinalam, a utilização de um *sistema gerenciador de banco de dados* (SGBD), isto é, de um conjunto de programas que torna possível a criação e a manutenção do banco de dados por seus usuários. Esse *software* atua como facilitador dos processos de definição, construção, manipulação e compartilhamento de banco de dados entre diversos usuários.

Elmasri e Navathe (2010, p.4) explicam que “a **definição** de um banco de dados implica especificar os tipos de dados, as estruturas e as restrições para os dados a serem armazenados em um banco de dados”. No caso da FN-Br, seu banco de dados é composto dos seguintes tipos de dados: *frames*, *ULs*, *construções* e *subcorpora*. Os *frames* constituem um conjunto de dados que apresentam uma definição, a qual compreende os Elementos de *Frame* que compõem o *frame*. Os Elementos de *Frame* são agrupados em tipos – nucleares, periféricos e extratemáticos. As *ULs* são os pareamentos de um lema a um *frame*. As *construções* constituem conjuntos de dados compostos de uma definição e da especificação de seus elementos. Os *subcorpora* são os conjuntos de sentenças que exemplificam, através da anotação em camadas, o comportamento de *ULs*, *frames* e *construções*.

“A **construção** de um banco de dados é o processo de armazenar os dados em alguma mídia apropriada controlada pelo SGBD”, assinalam Elmasri e Navathe (2010, p.4). No universo da FN-Br, o *software* Desktop é o nosso SGBD, na medida em que o banco de dados foi construído através desse *software*. A mídia de armazenamento utilizada são os servidores do projeto.

Os processos de **manipulação** compreendem funções, tais como “pesquisas em banco de dados para recuperar um dado específico, atualização do banco para refletir as mudanças no minimundo e gerar os relatórios dos dados” (ELMASRI & NAVATHE, 2010, p.4). Esse processo também é gerenciado pelo Desktop, visto que o *software* permite a criação de novos *frames* e construções, a anotação de sentenças, bem como a geração de relatórios de dados – os quais podem ser visualizados no *site* da FN-Br.

Por fim, como esclarecem Elmasri & Navathe (2010, p.4), “o **compartilhamento** permite aos múltiplos usuários e programas acessar, de forma concorrente, o banco de dados”. Na FN-Br, essa tarefa é realizada através dos relatórios gerados pelo Desktop, ou dos *data releases*. Assim, aquilo que se chama de “sistema de banco de dados” é a junção de um banco de dados ao *software* SGBD.

Considerando os pressupostos anteriormente apresentados, vejamos os problemas trazidos pela redundância que seria gerada caso anotássemos construções tratáveis como padrões de valência no Constructicon. Segundo os mesmos autores, a primeira questão a surgir é a de duplicação de esforços. Se uma mesma informação for anotada duas vezes em ambientes diferentes, quando for preciso realizar uma atualização, essa deverá ser feita, também, duas vezes. A segunda questão é o desperdício de espaço de armazenamento. Ao armazenar a mesma informação duas vezes, ocupamos um espaço que poderia ser utilizado para arquivar outras informações. Isso pode gerar problemas maiores para bancos de dados grandes. Por fim, há a questão da inconsistência de dados gerada pela anotação redundante. Ao se atualizar uma informação anotada duplamente em somente um dos ambientes utilizados para o armazenamento, uma incoerência entre os dados pode ser gerada se o mesmo não for feito no outro ambiente.

Contudo, conforme ponderam Elmasri e Navathe, em alguns casos é preciso utilizar a redundância controlada a fim de se aprimorar o desempenho em consultas. O SGBD deve atuar controlando essas redundâncias, para impedir inconsistências entre os arquivos.

Na implantação do Constructicon, optamos pela Não-Redundância para evitar os problemas acima apontados. Como ambos os projetos – FrameNet e Constructicon – se voltam a tarefas em PLN, seria ainda mais prejudicial ter

uma mesma informação anotada duas vezes, uma vez que o sistema poderia se confundir. Voltemos ao exemplo do tradutor automático, ao se deparar com uma sentença a ser traduzida, diante de uma informação dupla e paradoxal, por qual delas ele optaria para realizar a tradução? Por essa razão, a informação armazenada na FrameNet não pode constar novamente no Constructicon, e vice-versa.

É importante ressaltar que armazenar uma informação de maneira redundante não é sinônimo de anotar uma mesma sentença mais de uma vez. Como uma sentença é composta por mais de uma Unidade Lexical em potencial e, muitas vezes, instancia mais de uma construção, é possível que um mesmo segmento linguístico seja anotado mais de uma vez, porém, cada anotação trará informações diferentes para o banco, relacionando aquela sentença a ULs, construções e *frames* diferentes.

Ademais, considerando que a FrameNet foi elaborada com fundamentos construcionais (cf. discussão ao fim da sessão 2.1.2.), não se configura uma incongruência anotar uma construção nesse recurso na forma de um padrão de valência relacionado a uma UL.

Sabendo-se que construções gramaticais podem ser anotadas na FrameNet, uma questão surge. Se a FrameNet é capaz de abarcar construções gramaticais e, considerando que ela se baseia na Gramática das Construções, por que seria necessário desenvolver um Constructicon?

Parte dessa resposta concentra-se na argumentação elaborada por Fillmore (2008a) e aqui exposta na seção 2.3. Um dos pontos refere-se ao caso de multpalavras não tratáveis via padrão de valência, como, por exemplo, aquelas que requerem uma análise do contexto que as envolve. Outro ponto faz menção aos casos de anotação de texto corrido, modalidade que concentra um número maior de construções gramaticais. O autor aponta, ainda, a Construção WXDY como exemplo de construção cujos aspectos sintático-semânticos não podem ser inteiramente capturados por padrões de valência.

Além dos pontos levantados por Fillmore, as próprias construções da família *Para Infinitivo* evidenciam a necessidade de criação de um Constructicon, conforme a discussão apresentada no capítulo 3. Em suma, a necessidade de criação de um Constructicon repousa no fato de que diversos

esquemas sintáticos evocadores de *frames* não são processados e identificados em uma análise lexicográfica.

Uma vez que a implantação do Constructicon é necessária, e considerando que algumas construções podem ser tratadas via anotação lexicográfica, outra indagação emerge: como definir o que deve ser tratado via FrameNet e o que pode ser analisado via Constructicon? Em outras palavras, o que é definido como construção no âmbito do Constructicon? A próxima seção será permeada por essa discussão.

4.2 Reflexões sobre O Conceito de Construção

Mesmo com o desenvolvimento do Constructicon, vimos que determinadas construções gramaticais são anotadas na FrameNet. Logo, é preciso definir o que será construção no âmbito da FrameNet. É importante ressaltar que o que é concebido como construção em uma Gramática das Construções não o será, necessariamente, para o Constructicon, considerando que esse projeto tem objetivos outros além de descrever uma língua.

Assim, seguiremos três critérios para a definição de uma construção, isto é, para definir o que será tratado via Constructicon. Tais critérios emergem das observações relativas ao comportamento das construções da família estudada neste trabalho quando de seu tratamento pela FrameNet Brasil. Cabe ressaltar, ainda, que tais critérios devem ser vistos, seguindo os modelos cognitivistas da categorização. Portanto, o fato de uma construção em potencial encaixar-se no critério não deve ser visto como uma oposição binária, mas, sim, através da noção de grau de pertencimento, tal como ocorre com as categorias da Lógica Difusa.

Inicialmente, definiremos nosso objeto de análise, a construção em potencial, o esquema sintático a ser anotado. A primeira pergunta a ser levantada diz respeito à existência ou não de algum material linguístico lexicalmente especificado na construção. Como a anotação lexicográfica na FrameNet parte de uma Unidade Lexical, a qual se define como um lema pareado a um *frame*, é condição *sine qua non* a existência de algum item lexical específico para que esse tipo de anotação deva ocorrer. Assim, a

pergunta que devemos fazer acerca da construção em potencial é formalizável no critério 1:

Critério 1: Sendo X um material lexicalmente especificado, existe X na construção em potencial?

Se a resposta for negativa, sabemos que não será possível realizar uma anotação de tipo lexicográfico da construção e, portanto, ela deve ser anotada no Constructicon. Como todas as construções estudadas neste trabalho compartilham, ao menos, a preposição *para*, nenhuma delas satisfaz absolutamente esse critério. Porém, considerando-se que o Dativo com Infinitivo evoca o *frame* de Habilidade e não o de *Finalidade* – evocado por *para* – e que, afora a preposição, essa construção se constitui como um padrão muito genérico de estrutura argumental, ela se enquadra quase que totalmente nesse critério. O mesmo vale para as Construções de *Gap*. Assim, seguindo-se o primeiro critério, essas construções devem ser anotadas no Constructicon e não como ULs, em separado, sob pena de se perder em generalidade e adequação explicativa, conforme se discutiu no capítulo anterior.

Por outro lado, se a resposta for positiva, devemos partir para a segunda pergunta a ser levantada, a qual examina se a construção em potencial evoca ou não algum *frame*. Na anotação lexicográfica, as unidades básicas de análise são as ULs, pareamentos de um lema a um *frame*. Evocar um *frame* é, portanto, condição *sine qua non* para a unidade alvo. Se a construção em potencial não evoca *frame*, não pode ser abordada como UL e, destarte, exige um tratamento construcional. A pergunta a ser feita é formalizável no critério 2:

Critério 2: Sendo F um *frame* e X um material lexicalmente especificado, X evoca F?

Se a resposta for negativa, novamente, X deve ser encaminhado à descrição pelo Construction. Um tipo de construção que se adéqua a essa situação são os auxiliares de tempos compostos, como “ter” na sentença “tenho feito exercícios”. Há um material lexicalmente especificado, mas esse não é responsável por evocar nenhum *frame*. Trata-se de uma construção meramente formal.

Uma resposta positiva, porém, nos levaria à terceira e última pergunta, a qual aborda o fato de a construção em potencial evocar ou não o mesmo *frame* em padrões de valência distintos. Se uma unidade de material lexicalmente fixo for capaz de evocar um mesmo *frame* em padrões de valência diferentes significa que o *frame* foi evocado pela unidade alvo (também). Esse comportamento aponta para um tratamento lexicográfico, já que a anotação da UL, apenas, bastaria. Deste modo, a terceira pergunta é formalizável no critério 3:

Critério 3: Sendo F um *frame* e X um material lexicalmente especificado, X evoca F em outro padrão de valência?

A resposta negativa nos direciona ao tratamento construcional, ao passo que a positiva nos revela que X, na verdade, é uma UL e, portanto, pode ser tratado via padrão de valências.

A Construção Volitiva se aplicaria à terceira resposta negativa, uma vez que a mudança em seu esquema sintático – como exibido no capítulo 3 – muda o *frame* evocado. Enquanto “louco para” evoca o *frame* de volição extrema, “louco de” tem a própria UL “louco” como evocadora do *frame* de Propriedade_Mental. O mesmo ocorre com as construções modais e de tempo e aspecto.

A Construção Manipulativa, por outro lado, se adéqua à terceira resposta positiva, em que se constata que X é uma UL, visto que, no capítulo 3, pôde-se observar que a referida construção não tem seu *frame* alterado com a mudança em seu esquema sintático.

Doravante os anotadores da FrameNet poderão ser conduzidos pelos critérios acima expostos, a fim de se definir a qual ambiente de anotação

pertence cada construção. Tais critérios não haviam sido expostos ou esquematizados anteriormente pelos fundadores do Constructicon de Berkeley, embora tal discussão seja fundamental na constituição do *software*.

Do mesmo modo, navegando pelo Constructicon da língua inglesa, nos deparamos com uma série de falhas que precisam ser solucionadas. Veremos, a seguir, alguns exemplos dos problemas encontrados.

4.3 Inconsistências Presentes no Modelo do Constructicon

Conforme se discutiu na seção 3.2, a FrameNet e o Constructicon ainda não possuem ferramenta apta a mostrar o pareamento existente entre a forma de uma construção – seus ECs – e sua contraparte semântica – os EFs que constituem o *frame* que ela evoca. As duas partes do banco de dados não interagem. Se acessarmos, por exemplo, a Construção Modal Epistêmica como em “Dá para dormir na viagem?”, não encontraremos nenhuma forma de anotar os EFs do *frame* de Possibilidade, evocado pela construção. Também não encontraremos na descrição do *frame* de Possibilidade, nenhuma menção a essa construção como uma das entidades evocadoras desse *frame*. Se a construção também evoca *frame*, ela precisa, assim como a UL, estar presente na FrameNet. É preciso, portanto, fazer com que ambas as partes do banco de dados sejam integradas. Se as relações forem estabelecidas, mas não vinculadas, de nada servirão ao *software* responsável por processar tais informações.

O caso da Construção de Dativo com Infinitivo ilustra de maneira ainda mais contundente esse ponto. Se relembarmos a sentença utilizada em discussão na seção 3.2, “Eu comprei entrada pra assistir a uma determinada noite”, notaremos que o verbo “comprar”, sozinho, evoca o *frame* de Comércio_comprar, ao passo que a Construção de Dativo com Infinitivo é responsável por evocar o *frame* de Habilitação. Conseqüentemente, temos que há *frames* que são evocados pelos verbos e outros que são evocados pela construção.

Embora tais *frames* tenham origens diferentes – nessa situação em particular, os constituintes evocadores são um verbo e uma construção –, eles interagem. No caso da Construção de Dativo com Infinitivo, essa interação é

feita na medida em que o *frame* verbal é responsável por apontar de que modo o Recurso foi obtido. Tomando-se, novamente, a sentença acima, sabemos que o Recurso obtido “entrada” (evocado pela construção) foi garantido através de uma compra (evocado pelo verbo). Em outro exemplo também trabalhado na seção 3.2, “Então eles me davam aquele dinheiro para eu ir no bonde normal”, a interação entre os *frames* também existe; ela indica que o Recurso “aquele dinheiro” fora dado, já que o *frame* evocado pelo verbo é o de Dar.

Vale ressaltar, ainda, que a evocação de dois *frames* diferentes para uma mesma sentença não constitui uma redundância. Ainda que tal sentença venha a ser anotada no ambiente da FrameNet lexicográfica, e também no Constructicon, o que ocorre é a atribuição de informações diferentes sobre o mesmo dado, o que, portanto, não configura uma redundância.

Logo, a análise que aqui se propõe, a qual visa a anotar um *frame* para a construção, e outro para o verbo ou para o adjetivo, coincide com a proposta de Goldberg de que *frames* verbais e construcionais interagem no construto – conforme discussão da seção 2.2.1.1.

Ademais, o Constructicon não possui ferramenta similar ao *FrameGrapher* usado na FrameNet, nem mesmo uma descrição das relações existentes entre as construções. Como não há um estudo dessas relações, encontramos lacunas ainda maiores. Analise a seguinte situação: se observarmos a descrição da Construção *Wherewithal*, encontraremos menção de que essa construção herda da Construção da Oração Relativa. No entanto, essa última não foi descrita no Constructicon; no recurso, ela não existe. Tampouco há uma definição do que seja uma relação de herança no Constructicon. A importância da definição e aprimoramento das relações entre *frames* e entre construções é clara e tal reflexão nos leva a outros questionamentos, como o seguinte: é possível colocar as relações entre construções em paralelo com as relações entre *frames*? Para justificar a questão, é preciso que passemos por alguns pontos.

Goldberg (1995) descreve o Princípio da Motivação Maximizada dizendo que, se uma construção se relaciona a outra sintaticamente, então essa construção também se relaciona semanticamente à outra, de modo que a motivação foi maximizada. Se no decorrer dessa investigação for possível pôr em paralelo as relações construcionais com as relações entre *frames*, o

Princípio se comprovaria. Como a construção também é um elemento capaz de evocar um *frame*, duas construções relacionadas pela forma (o esquema da família Para Infinitivo, por exemplo), teriam os *frames* por elas evocados também relacionados, o que acarretaria no paralelismo entre as relações.

Na descrição das relações construcionais, é necessário aliar uma descrição mais detalhada capaz de captar mais generalizações, conforme Goldberg propõe, ao trabalho feito por Fillmore, mais formal, já que, é preciso lembrar que a FrameNet é, antes de tudo, um recurso empregado em tarefas de PLN. É preciso que as descrições das relações sejam aprimoradas a fim de se fornecer, verdadeiramente, uma ferramenta capaz de prover informações completas sobre *frames* e construções

Claramente, intervenções no modelo de sistema de dados do Constructicon que lhe deem condições de dar conta desse conjunto de necessidades requerem um arcabouço de reflexões teóricas que extrapolam o escopo desta dissertação. Assim, nossa proposta circunscreve-se ao estabelecimento dos critérios necessários à definição de uma construção no âmbito da FrameNet, tanto no que diz respeito ao fato de que os ECs devam ser definidos com base em suas propriedades formais e, posteriormente, mapeados aos EFs que evocam, quanto no que concerne às perguntas que definem se dado material linguístico deve ou não ser tratado no Constructicon. Tais critérios não haviam sido postulados na proposta do Constructicon de Berkeley, o que contribuiu para que se gerassem as diversas lacunas supracitadas. Estabelecer tais critérios é importante para que o processo de anotação seja formalizado e ganhe contornos rígidos que impeçam problemas, tais como o de redundância de dados e a incerteza relativa ao ambiente adequado para se anotar uma dada unidade alvo.

Em trabalhos futuros, pretendemos endereçar tais questões, começando pela revisão das relações entre *frames* e entre construções na FrameNet.

5 CONCLUSÕES

Ao longo desta dissertação buscou-se fundamentar as discussões teórico-metodológicas para a implementação do Constructicon para o português brasileiro. Para tanto, onze construções da família *Para Infinitivo* foram utilizadas a fim de se estabelecer critérios para um tratamento construcional.

Tais critérios foram constituídos tendo-se em vista o direcionamento que tem o Constructicon a tarefas de PLN, o que nos faz refletir sobre categorias outras além daquelas que concernem as descrições linguísticas. A questão da Não-Redundância em banco de dados, por exemplo, aponta para o fato de que anotar uma construção no ambiente lexicográfico e, ao mesmo tempo, no construcional é prejudicial para o Constructicon, caso as informações geradas a partir de cada anotação sejam idênticas, e.g. caso os EFs evocados pelas ULs e pelas construções sejam os mesmos.

Outras questões norteadoras de tais critérios dizem respeito a incoerências em se anotar, por exemplo, uma construção que não evoca *frames* na FrameNet, ou, ao contrário, anotar no Constructicon uma construção que evoca o mesmo *frame* em padrões de valência diferentes. Assim, os critérios estabelecidos são os seguintes:

Critério 1: Sendo X um material lexicalmente especificado, existe X na construção em potencial?

Critério 2: Sendo F um *frame* e X um material lexicalmente especificado, X evoca F?

Critério 3: Sendo F um *frame* e X um material lexicalmente especificado, X evoca F em outro padrão de valência?

Doravante, ao se anotar uma construção, tais critérios deverão ser seguidos a fim de se definir o que seja uma construção no âmbito da FrameNet.

Além de estabelecer tais critérios, os quais não haviam sido formalizados nem discutidos pelos criadores do Constructicon de Berkeley, o que dificultava quando da decisão do ambiente adequado a se anotar uma dada unidade, a presente dissertação apontou falhas no empreendimento do primeiro Constructicon. Essas recobrem problemas de origens diversas, como, por exemplo, a não-interação entre os bancos de dados da FrameNet e do Constructicon, problema grave já que ambos os recursos se voltam a pesquisas em PLN.

Em conclusão, esta dissertação propõe os seguintes avanços relativos à sua contribuição teórica:

- (i) Ao se estabelecer os critérios que definem uma construção no âmbito da FrameNet, formaliza-se uma definição necessária aos processos de anotação. Até então, a definição de construção, emprestada das vertentes em Gramática das Construções, não era adequada ao Constructicon, uma vez que apresentava ambiguidades, visíveis, sobretudo, quando da anotação de construções.
- (ii) Ao se apontarem as falhas observadas no Constructicon, abre-se caminho para, em uma investigação futura, se buscar solucionar tais problemas, seja no domínio teórico, ou na composição do Constructicon como ferramenta de PLN que é.

REFERÊNCIAS BIBLIOGRÁFICAS

BICK, E. **The Parsing System PALAVRAS**: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Arhus, Arhus University, 2000.

BOAS, H. C. Cognitive Construction Grammar. In: HOFFMANN, T.; TROUSDALE, G. **The Oxford Handbook of Construction Grammar**. Oxford: Oxford University Press, 2010.

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

CROFT, W. **Syntactic Categories and Grammatical Relations**. Chicago: University of Chicago Press. 1991.

_____. **Aspectual and causal structure in event representations**. Manuscript. University of New Mexico. 2005.

ELMASRI, R.; NAVACHE, S. B. **Sistemas de banco de dados**. 6ª São Paulo: Pearson, 2010.

FERRARI, L. **Introdução à Linguística Cognitiva**. São Paulo: Contexto, 2011.

FILLMORE, C. J. The case for case. In: BACH, E. & HARMS, R. (eds) **Universals in Linguistic Theory**. New York: Holt, Rinehart and Winston, 1968.

_____. The grammar of hitting and breaking. In: JACOBS, R. A.; ROSENBAUM, P. S. (eds). **Readings in English Transformational Grammar**. Waltham MA: Ginn and Company, 1970.

_____. The case for case reopened. In: COLE, P.; SADDOCK, J. (eds) **Grammatical Relations**. New York, Academic Press, p.59-81, 1977.

_____. Frame Semantics. In: THE LINGUISTIC SOCIETY OF KOREA (org.). **Linguistics in the morning calm**. Seoul: Hanshin, 1982.

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

_____. The Mechanisms of “Construction Grammar. **Berkeley Linguistic Society**, 14, 35–55, 1988.

_____. Border Conflicts: FrameNet Meets Construction Grammar. In: **EURALEX,13, 2008, Barcelona. Anais...** Barcelona: Universitat Barcelona Fabra, 2008a.

_____. Wherewithal: eine verborgene Konstruktion. In: STEFANOWITSCH, A.; KERSTIN, F. (eds.). **Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik**. Tübingen: Stauffenburg, 279-283, 2008b.

FILLMORE, C. J.; KAY, P. Grammatical Constructions and Linguistic Generalizations: the What's X Doing Y Construction. **Language**, vol. 75, n° 1, 1999.

FILLMORE, C.; JOHNSON, C.; PETRUCK, M. **Background to FrameNet**. In: International Journal of Lexicography. Vol 16 n° 3. Oxford University Press, 2003.

FILLMORE, C. J.; LEE-GOLDMAN R.; RODHES, R. The FrameNet Constructicon. In: BOAS, H. C.; SAG, I. **Sign-Based Construction Grammar**. Chicago: CSLI, 2012.

FRIED, M.; OSTMAN, J-O. Construction Grammar: a thumbnail sketch. In: _____. Construction Grammar in a Cross-Linguistic Perspective. Amsterdam: John Benjamins, 2004.

GEERAERTS, D.; CUYCKENS, H. "Introducing Cognitive Linguistics", IN: GEERAERTS, D.; CUYCKENS, H. (eds). The Oxford Handbook of Cognitive Linguistics. Oxford/New York: Oxford University Press, 2007.

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.

_____. **Verbs, constructions, and semantic frames**, in M. Rappaport Hovav, E. Doron, I. Sichel (Eds), Oxford University Press, New York, NY. 2010.

LAKOFF, G. **Women, Fire and Dangerous Things: What categories reveal about the mind**. Chicago: University of Chicago Press, 1987.

LEE-GOLDMAN, R. **Context in Constructions**. Ph.D. Dissertation in Linguistics. University of California, Berkeley, 2011.

OTHERO, G. A.; MENUZZI, S. M. **Lingüística Computacional: teoria & pratica**. São Paulo: Parábola Editorial, 2005.

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

PETRUCK, M.R.L.; MELO, G. Precedes: A Semantic Relation in FrameNet. In: **Proceedings of the Workshop on Language Resources for Public Security Applications** at the 8th Conference on International Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, pp.45-49, 2012.

RUPPENHOFER, J.; ELLSWORTH, M.; PETRUCK, M.; JOHNSON, C.; SCHEFFCZYK, J. **FrameNet II: Extended Theory and Practice**. Versão 14 set. 2010. Disponível em: <http://framenet.icsi.berkeley.edu/> Acesso em 4 de junho de 2012.

SAG, I.; BOAS, H. C.; KAY, P. **Introducing Sign-Based Construction Grammar**. In: BOAS, H. C.; SAG, I. **Sign-Based Construction Grammar**. Chicago: CSLI, 2012.

SALOMÃO, M. M. M. **Polysemy, aspect and modality in Brazilian Portuguese: the case for a cognitive explanation of grammar**. 1990. 295 f. Tese de Ph.D. em Linguística, University of California at Berkeley, Berkeley, 1990.

_____. **Lanterna na proa: sobre a tradição recente nos estudos de linguística. Gragoatá**. n.23, p.52, Niterói, Editora da UFJF, 2007a.

_____. **The Metaphoric Grounding of Grammar: the Modal Construction with 'give' in Brazilian Portuguese**. In: **The 33rd Annual Meeting of The Berkeley Linguistics Society**. University of California, Berkeley. 2007b.

_____. **FrameNet Brasil: um trabalho em progresso. Calidoscópico**, São Leopoldo: UNISINOS, vol. 7 n. 3, p. 171-182, set/dez 2009a.

_____. **Tudo certo como dois e dois são cinco: Todas as construções de uma língua**. In: MIRANDA, N. S. & SALOMÃO, M. M. M. (Org.). **Construções do Português do Brasil: da gramática ao discurso**. Belo Horizonte: UFMG, 2009b.

SAMPAIO, T. F. **A Família de Construções de Argumento Cindido no Português do Brasil**. Tese de Doutorado em Linguística. Universidade Federal de Juiz de Fora, Juiz de Fora, 2010.

SAUSSURE, F. **Curso de Lingüística Geral**. 29ª ed. São Paulo. Cultrix: 2008 [1916].

SCHIFFRIN, D. **Discourse Markers**. Cambridge, UK: Cambridge University Press, 1987.

SILVA, A. B. C.; ESTRADA, C. C. D. **Coesão e coerência em redações do enem sob a perspectiva da FrameNet Brasil**. Monografia de pós-graduação em Ensino de Língua Portuguesa. UFJF, Juiz de Fora, 2012.

SILVA, A. B. L.; TORRENT, T. T. **Construções em Contraste: uma análise da Construção Habilitativa em Para Infinitivo do Português do Brasil e da Construção Wherewithal do Inglês**. In: **Anais do XI Encontro de Linguística de Corpus (ELC 2012)**. São Carlos: USP, 2012.

TORRENT, T. T. **A Rede de Construções em Para (SN) Infinitivo**: uma abordagem centrada no uso para as relações de herança e mudança construcionais. 2009. 166 f. Tese de Doutorado em Linguística. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

_____. **The Construction Network Reconfiguration Hypothesis**. In: Rodrigues, A.; Braga, Maria L.; Torrent, T. T. (Org.). LETRAS & LETRAS. Uberlândia: UFU, Volume 27 - Número 1 - jan./jun.. 2011.

_____. **Frames e Construções**: a anotação de esquemas construcionais na FrameNet Brasil. In: *Anais do I Congresso Internacional de Estudos do Léxico*. Salvador: UFBA, 2012.

_____. **On The Relation Between Inheritance and Change**: The Constructional Convergence and The Construction network Reconfiguration Hypotheses. (no prelo).

VIEIRA, R.; LIMA, V. L. S. **Linguística computacional: princípios e aplicações**. In: NEDEL, L. (Ed.) **IX Escola de Informática da SBC-Sul**. Passo Fundo, Maringá, São José: SBC-Sul, 2001.

_____. **Linguística Computacional: Princípios e Aplicações**. Porto Alegre: Pucrs, 2011.