

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

Ludmila Meireles Lage

**MODELAGEM LINGUÍSTICO-COMPUTACIONAL DAS RELAÇÕES ENTRE
CONSTRUÇÕES E FRAMES NO CONSTRUCTICON DA FRAMENET BRASIL**

Juiz de Fora
Janeiro de 2018

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

Ludmila Meireles Lage

**MODELAGEM LINGUÍSTICO-COMPUTACIONAL DAS RELAÇÕES ENTRE
CONSTRUÇÕES E FRAMES NO CONSTRUCTICON DA FRAMENET BRASIL**

Tese de Doutorado 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 Doutor em Linguística.

Orientador: Professor Doutor Tiago Timponi Torrent

Juiz de Fora
Janeiro de 2018

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Lage, Ludmila Meireles.

Modelagem linguístico-computacional das relações entre construções e frames no Constructicon da FrameNet Brasil / Ludmila Meireles Lage. -- 2018.

130 f.

Orientador: Tiago Timponi Torrent

Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2018.

1. Constructicon. 2. FrameNet. 3. Relações. 4. Gramática das Construções. 5. Semântica de Frames. I. Torrent, Tiago Timponi, orient. II. Título.

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

Ludmila Meireles Lage

**MODELAGEM LINGUÍSTICO-COMPUTACIONAL DAS RELAÇÕES ENTRE
CONSTRUÇÕES E FRAMES NO CONSTRUCTICON DA FRAMENET BRASIL**

Tese de Doutorado 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 Doutor em Linguística.

Aprovada por:

Presidente, Prof. Dr. Tiago Timponi Torrent – UFJF

Profa. Dra. Flávia Bezerra Hirata-Vale – UFSCAR

Prof. Dr. Diogo Oliveira Ramires Pinheiro – UFRJ

Profa. Dra. Maria Margarida Martins Salomão – UFJF

Prof. Dr. Luiz Fernando Matos Rocha – UFJF

Juiz de Fora
Janeiro de 2018

*Para meu amor, Victor,
que sonhou junto comigo*

AGRADECIMENTOS

Agradecer é um ato de prestar honras àquele que realizou benfeitorias. Logo, fazê-lo é justo e desejável, sobretudo quando tantas situações prósperas nos saúdam pelo caminho. Assim, agradeço, primeiramente, à instituição que já chamo de casa, a Universidade Federal de Juiz de Fora. Agradeço à UFJF, que resiste na luta pelo ensino de qualidade e gratuito. Estendo minha gratidão ao Programa de Pós-Graduação em Linguística, que sempre se esforçou em apoiar a pesquisa nas variadas veredas a que ela conduz.

À FAPEMIG, agradeço por todo o suporte no decorrer do doutorado, bem como ao programa de Doutorado Sanduíche CAPES/STINT, que permitiu à minha investigação ampliar seu escopo, ao alçar vôos mais altos.

Ao meu orientador, Tiago Torrent, por quem tenho grande admiração, direciono toda a minha gratidão. Entre tantos outros fatores, agradeço pela atenção, pela dedicação e, principalmente, pelo bom humor, que tornou reuniões e orientações muito mais agradáveis. Obrigada por nos inspirar a almejar sempre mais, sendo o grande incentivador dos vôos alçados – e daqueles que ainda alçaremos.

Ao meu estimado orientador na Suécia, Benjamin Lyngfelt, agradeço infinitamente pela acolhida na escura Escandinávia invernal. Sou grata por abrir-me caminhos e mente, pelas conversas e pela amizade que ficou.

Depois de quase nove anos, a FrameNet Brasil acabou se tornando também sinônimo de amizades feitas. Dessa forma, agradeço aos queridos colegas que comigo dividiram seus questionamentos, descobertas e sucesso. Em especial, sou grata a Maucha e a Adrieli. A Maucha, agradeço por sonhar, sofrer e comemorar junto comigo as variadas etapas dessa jornada na pós-graduação. Seguiremos juntas! A Adrieli, agradeço por ter se tornado minha família em Gotemburgo. Obrigada pela amizade e por tudo que compartilhamos. Foi fundamental ter vivido parte da Suécia com você!

Agradeço a Natália e Cláudia, que sempre incentivaram meu desenvolvimento e o de minha pesquisa de doutorado. Minha profunda gratidão pelo apoio ao Doutorado Sanduíche!

Aos queridos amigos da vida, agradeço imensamente o apoio e incentivo por tudo que vivi nos últimos anos, desde as muitas vezes em que estive ausente, escrevendo, até quando decidi partir para o Sanduíche na Suécia. Vocês foram maravilhosos!

E minha mais profunda gratidão à minha amada família, Victor, Auxiliadora e Maurício. Obrigada por serem meu alicerce, sem o qual eu não poderia trilhar este caminho. Aos meus pais, agradeço pelo encorajamento e estímulo. Se hoje escrevo esta tese, é porque guardei o que sempre ouvi de vocês: que os estudos deveriam ser prioridade em minha vida. Obrigada pelos valiosos ensinamentos sobre o valor do conhecimento! Ao meu amor, agradeço a compreensão, a paciência e o suporte. Obrigada por ser meu maior incentivador, por sonhar meus sonhos, e por me fazer acreditar em mim mesma. Você me faz voar.

RESUMO

Esta tese apresenta as discussões teórico-metodológicas que embasaram a modelagem linguístico-computacional das relações entre construções e entre construções e frames no Constructicon da FrameNet Brasil. Após a fase de implementação de tal recurso, desenvolvido a fim de explicar os fenômenos não capturados pelas análises lexicográficas proporcionadas pela FrameNet, foi necessário transformá-lo em uma rede. Uma vez que construções são concebidas como construtos cognitivos que participam em redes relacionais, foi modelada a relação de Herança (KAY & FILLMORE, 1999) para estruturar a rede de construções. Outro importante avanço implementado foi a relação de Evocação, que captura os casos em que uma construção evoca um frame, conectando-os de modo a evidenciar a relação entre eles. As construções e os frames, dois dos marcos teóricos mais importantes da Linguística Cognitiva, têm sido, de fato, foco de estudos quanto às relações que estabelecem entre si. Ademais, dado que o Constructicon foi desenvolvido em paralelo com a FrameNet, teria sido um desperdício não conectar ambos os bancos de dados. Contudo, alguns aspectos apresentados pelas construções sobrepujam as generalizações capturadas pela herança e pela importação semântica representada em termos de frames. Além disso, visto que o Constructicon da FrameNet Brasil se volta sobretudo a tarefas em tecnologia da linguagem, era preciso que o recurso proporcionasse informações que fossem legíveis não apenas a leitores humanos, mas também a máquinas. Dessa forma, através da modelagem das construções Aspectual Inceptiva (SIGILIANO, 2012, 2013) e Dativo com Infinitivo (TORRENT, 2009; LAVIOLA 2015), observou-se a necessidade de se adicionar ao Editor de Restrições mecanismos para dar conta de tais aspectos. De tal modo, doravante é possível registrar as unidades lexicais (ULs) que podem preencher um elemento da construção. Essa propriedade pode ser implementada em três níveis diferentes: caso apenas ULs específicas possam preencher um determinado slot; caso todas as ULs que evocam um frame possam preencher o dado slot; ou caso uma família de frames seja aceita em um slot. Assim sendo, através da implementação das relações descritas, este trabalho contribui para o progresso do recurso construcional, bem como para exibir os mecanismos necessários para diversas aplicações computacionais, ao mesmo tempo em que operacionaliza, no domínio computacional, a interdependência entre frames e construções há tanto tempo pontuada pelos estudos descritivos em Linguística Cognitiva.

Palavras-chave: Constructicon; FrameNet; Relações; Gramática das Construções; Semântica de Frames.

ABSTRACT

This work presents the theoretical-methodological discussions that support the linguistic-computational modeling of the relations between constructions and between constructions and frames in the FrameNet Brasil Constructicon. After implementing the resource, developed to explain the phenomena not captured by the lexicographic analyzes provided by FrameNet, it was necessary to turn it into a network. Since constructions are conceived as cognitive constructs that participate in relational networks, the Inheritance relation (KAY & FILLMORE, 1999) was modeled for structuring a constructions network. Another important advance implemented was the Evokes relation, which captures the cases in which a construction evokes a frame, connecting them to put in evidence the relation between them. Constructions and frames, two of the most important theoretical frameworks of Cognitive Linguistics, have been indeed the focus of studies on the relations they establish with each other. In addition, since the Constructicon was developed in parallel with FrameNet, it would have been a waste to not connect both databases. However, some aspects of constructions transcend the generalizations captured by inheritance and by the semantic import represented in terms of frames. Moreover, since FrameNet Brasil Constructicon focuses primarily on language technology tasks, the resource needed to provide information that would be readable not only to human readers but also to machines. Thus, through the modeling of Inceptive Aspectual (SIGILIANO, 2012, 2013) and Dative with Infinitive (TORRENT, 2009; LAVIOLA 2015) constructions, it was observed the need to add some mechanisms to the Constraint Editor to account for these aspects. Therefore, henceforth it is possible to register the lexical units (LUs) that can fill a construction element. This property can be implemented in three different levels: if only specific LUs can fill a given slot; if all LUs evoking a frame can fill a given slot; or if a frame family is accepted in a slot. Hence, through the implementation of the relations described, this work contributes to the progress of the constructional resource, as well as to display the mechanism applicable to several computational applications, at the same time it operates, in the computational domain, the interdependence between frames and constructions that have been long punctuated by the descriptive studies in Cognitive Linguistics.

Keywords: Constructicon; FrameNet; Relations; Construction Grammar; Frame Semantics.

SUMÁRIO

INTRODUÇÃO	10
1 GRAMÁTICAS DE CONSTRUÇÕES: FUNDAMENTOS	14
1.1 Gramática das Construções de Berkeley	16
1.2 Gramática Cognitivista das Construções.....	25
1.3 Gramática das Construções Corporificada.....	30
2 LINGÜÍSTICA COMPUTACIONAL: FUNDAMENTOS E IMPLEMENTAÇÕES COGNITIVAMENTE INSPIRADAS	35
2.1 Fundamentos da Linguística Computacional	35
2.2 Níveis de Análise	37
2.3 Implementações Computacionais Cognitivamente Inspiradas	39
2.3.1 <i>Lexica Baseados em Frames</i>	39
2.3.2 <i>Constructica</i>	50
2.3.2.1 Fundamentos	50
2.3.2.2 Constructicon da FrameNet Brasil.....	55
3 RELAÇÕES ENTRE CONSTRUÇÕES E ENTRE FRAMES E CONSTRUÇÕES	63
3.1 Redes Construcionais: Herança Completa versus Herança Normal	63
3.2 Relações entre Frames e Construções.....	72
3.3 Implementações Computacionais.....	75
3.3.1 <i>O ECG Analyzer</i>	76
3.3.2 <i>O Constructicon de Berkeley</i>	79
3.3.3 <i>O Constructicon do Sueco</i>	82
4 METODOLOGIA	86
5 MODELAGEM DAS RELAÇÕES ENTRE CONSTRUÇÕES E ENTRE FRAMES E CONSTRUÇÕES NO CONSTRUCTICON DO PORTUGUÊS BRASILEIRO	89
5.1 Relações entre Construções	89
5.1.1 <i>Herança Construcional</i>	90
5.1.2 <i>Implementação de Exemplo: Construções Relativas</i>	91
5.2 Relações entre Frames e Construções	101
5.2.1 <i>Evocação</i>	102
5.2.2 <i>Implementação de Exemplo: Construções de Dativo com Infinitivo</i>	104
5.2.3 <i>Restrição Semântica de Preenchimento de Slot</i>	108
5.2.4 <i>Implementação de Exemplo: Construções Aspectuais Inceptivas</i> ..	108
6 CONCLUSÕES	119
REFERÊNCIAS BIBLIOGRÁFICAS	121

INTRODUÇÃO

Esta tese tem por objetivo apresentar as discussões teórico-metodológicas que embasaram a modelagem linguístico-computacional das relações entre construções e entre frames e construções no Constructicon da FrameNet Brasil. Tal recurso, um repertório online de construções do português brasileiro (PB), se desenvolve em paralelo ao Lexicon da mencionada FrameNet Brasil, recurso, também disponível online, que descreve unidades lexicais do PB com base na Semântica de Frames.

Quando da implementação do Constructicon, foi imprescindível estabelecer critérios definidores dos materiais linguísticos que seriam analisados via anotação construcional, e aqueles que poderiam ser tratados por meio da anotação lexicográfica, uma vez que algumas construções eram plenamente descritas pelos padrões de valência (LAGE, 2013).

Após a fase inicial, desenvolvida ao longo da pesquisa de mestrado, notou-se que, diferentemente da estrutura de rede que apresentava a FrameNet Brasil, cujos frames se conectavam por diversas relações, o Constructicon necessitava se tornar uma rede do mesmo modo, a fim de permitir generalizações e garantir economia ao mecanismo. Esta foi, portanto, a primeira motivação para o desenvolvimento desta tese: estruturar o Constructicon em uma rede hierárquica de construções. Para tanto, seria necessário investigar qual (ou quais) relação (ões) estabelecer a fim de proporcionar descrições satisfatórias do ponto de vista linguístico bem como do computacional. Visto que o Constructicon não tem os mesmos objetivos de uma Gramática das Construções (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012), outros elementos passam a ser importantes – tais como redundância do sistema, sua complementaridade em relação ao Lexicon e suas aplicações.

Para além dessa tarefa, uma outra e mais profunda questão se impôs. Ao investigarmos variadas construções, verificaram-se diversos padrões construcionais que seriam plenamente descritos se desfrutassem de um robusto aparato semântico. Conforme visto, o Constructicon se desenvolve em paralelo ao já estabelecido Lexicon da FrameNet Brasil e, por conseguinte, à toda modelagem semântica que o embasa. Além disso, conforme enunciado desde os primeiros trabalhos em Gramática das Construções, itens lexicais também

são construções (FILLMORE, KAY & O'CONNOR, 1988, p.501), porém, menos esquemáticas. Logo, a seguinte pergunta de pesquisa emergiu:

Sendo (a) a FrameNet Brasil um modelo semântico já estabelecido e (b) os itens lexicais um subtipo de construção, seria tal recurso adequado para modelar a contraparte semântica das construções descritas no Constructicon?

Boas (2017, p.571) alega que os dados contidos nas FrameNets de todos os idiomas podem ser considerados como um depósito de informações semânticas baseadas em frames de múltipla aplicabilidade para pesquisas em Gramática das Construções. Afirmo, ainda, que, já que construções são pareamentos de uma forma a um significado, o polo do significado pode ser representado pelos frames. O pesquisador lembra, também, que Fillmore (1988) foi quem primeiro demonstrou a integração sistemática da Semântica de Frames à análise construcional.

Lee-Goldman & Petruck (no prelo) recordam a história intelectual compartilhada pela Semântica de Frames e pela Gramática das Construções referindo-se ao visionário artigo “The Case for Case” de Fillmore (1968). Produtos das chamadas “teorias irmãs”, já que comungam da mesma gênese, frames e construções não necessariamente têm sido tratados de forma interdependente. Há implementações descritivas que unem ambas as teorias, tais como as listadas por Boas (2017): Kay & Fillmore (1999), Michaelis & Ruppenhofer (2001), Boas (2003), Hasegawa et al. (2010) entre outras. Entretanto, nenhuma implementação computacional o faz efetivamente – ou não se lida com frames como aporte semântico – caso da Embodied Construction Grammar –, ou não se estabelece uma relação estrutural entre os aparatos teóricos – caso dos Constructicons de Berkeley e do Sueco.

Portanto, esta é uma questão não só de cunho metodológico-computacional, como à primeira vista se apresenta, mas também de contribuição teórico-linguística, já que, sendo positiva a resposta à questão desta tese, torna-se viabilizada a união das referidas teorias no campo descritivo e computacional.

De tal modo, buscou-se responder a tal questão através da modelagem das relações entre construções e frames, assumindo-se as seguintes hipóteses:

- i. A FrameNet Brasil e, por extensão, as demais framenets, prestam-se a modelar a contraparte semântica das construções descritas computacionalmente em um Constructicon.**
- ii. O aparato de software existente para a FrameNet Brasil pode ser utilizado para modelar relações entre construções em um Constructicon.**

Com o objetivo de testar as hipóteses supra enunciadas, este trabalho encontra-se dividido em seis capítulos. Primeiramente, é apresentada a Gramática das Construções, em especial em três de suas vertentes, quais sejam, a Gramática das Construções de Berkeley, a Gramática Cognitivista das Construções e a Gramática das Construções Corporificada. O objetivo deste capítulo é apresentar os fundamentos dos três modelos construcionistas que, de alguma forma, desempenharam um papel na perseguição do objetivo desta tese, além de fornecer o conhecimento terminológico fundamental para as abordagens em Gramática das Construções.

O capítulo 2 aborda a Linguística Computacional quanto a seus principais fundamentos, além de expor implementações computacionais que se valem da Linguística Cognitiva.

Por sua vez, o capítulo 3 discute as versões da Gramática das Construções, levantadas no primeiro capítulo, pelo viés da herança construcional e das implementações que lidam com frames e construções.

No capítulo 4, é exposta a metodologia de trabalho empregada no Constructicon e o capítulo 5 traz as análises e discussões que fundamentam a modelagem linguístico-computacional das relações entre construções e entre frames e construções propostas nesta tese, no âmbito da FrameNet Brasil.

Por fim, o capítulo 6 conclui este trabalho, apontando as contribuições trazidas por esta tese, quais sejam:

- i. Converter o Constructicon da FrameNet Brasil em uma rede construcional sustentada pela relação de herança;

- ii. Interligar de forma sistemática as bases de dados lexical e construcional;
- iii. Estabelecer um *locus* para a união das teorias frames e construções, revelando a complementaridade de tais construtos mesmo em recursos computacionais altamente sistematizados.

1 GRAMÁTICAS DE CONSTRUÇÕES: FUNDAMENTOS

Fried e Ostman (2004, p.12), ao caracterizar os propósitos de uma gramática das construções, afirmam que essa não difere de qualquer outra teoria de gramática na medida em que também busca encontrar a melhor maneira de representar a relação entre estrutura, significado e uso. Entretanto, diferentemente de outras teorias, essa gramática compreende forma e função como elementos inseparáveis. Assim sendo, ela assume a construção gramatical como unidade básica de análise e representação. Por “construção”, entende-se um pareamento de uma forma a um significado – embora haja alguma divergência quanto ao caráter absoluto dessa concepção¹.

A base conceptual da Gramática das Construções, conforme apontam os autores, sustenta-se por três hipóteses:

- (i) falantes apoiam-se em padrões de forma e significado relativamente complexos – construções – para construir expressões linguísticas;
- (ii) expressões linguísticas refletem os efeitos da interação entre construções e o material linguístico, tais como as palavras, que ocorrem nelas; e
- (iii) construções são organizadas em redes de padrões sobrepostos, relacionados por meio de propriedades compartilhadas. (FRIED & OSTMAN, 2004, p.12)^{2,3}

Deste modo, visto que todo padrão de forma e significado corresponde a uma construção, unidades diferentes – como palavras e orações – não são tratadas de maneira diferenciada. Em outras palavras, não é atribuído papel central a nenhuma unidade linguística ou padrão gramatical. Neste sentido, Kay e Fillmore (1999, p.1-2) ressaltam que adotar a abordagem construcional é assumir que padrões de diferentes níveis são igualmente objetos a serem analisados por essa gramática – contemplando-se desde as formas mais regulares até as expressões idiomáticas. Assim, essa gramática propõe-se a representar, por meio de um sistema explícito, todas as construções sem perda de generalização e, ao mesmo tempo, de forma econômica.

¹ Sobre o referido desacordo, discutiremos nas próximas seções.

² Todas as traduções contidas nesta tese são de responsabilidade da autora. Inserem-se em nota de rodapé os trechos originais.

³ (i) speakers rely on relatively complex meaning-form patterns – constructions – for building linguistic expressions; (ii) linguistic expressions reflect the effects of interaction between constructions and the linguistic material, such as words, which occur in them; and (iii) constructions are organized into networks of overlapping patterns related through shared properties.

Kay e Fillmore (1999, p.2) ressaltam que a presente proposta é a de uma gramática explícita, não derivacional, baseada na noção de construção gramatical como uma associação convencional de forma linguística e conteúdo. Na Gramática das Construções, não se aposta em regras de combinação entre unidades linguísticas para produção de sentido, isto é, a associação entre duas ou mais formas “raramente resulta em uma simples concatenação de sentidos que essas formas podem ter isoladamente” (FRIED & OSTMAN, 2004, p.12)⁴. Salomão (2009a, p.39) exemplifica tal afirmativa por meio dos lexemas *carcereiro* e *prisioneiro*, ressaltando que ambos são formados pela associação do sufixo {+eiro} a duas raízes nominais sinônimas (*prisão* e *cárcere*) e, no entanto, eles resultam em papéis opostos do frame de *Prisão*.⁵ Em uma análise composicional, o significado das duas palavras teria de ser o mesmo, o que ignoraria as complexidades dos elementos constituintes de tais lexemas, como o caráter polissêmico do sufixo em questão. Em suma, a construção é o fator determinante nessa análise.⁶

Conforme salientam Fried e Ostman (2004, p.14-15), os papéis semânticos assumem posição central na Gramática das Construções. Os autores recordam a sugestão de Kay e Fillmore (1999) de que a gramática pode ser compreendida como uma rede de associações entre papéis sintáticos (funções gramaticais), papéis textuais (estrutura de informação) e papéis situacionais (aqueles dependentes do frame em questão – como *Comprador* e *Vendedor*, na cena de *Transação_comercial* descrita por Fillmore (1985).

Fried e Ostman (2004, p.18) explicam que construções são signos simbólicos que representam os blocos básicos de análise linguística – desde um morfema até um texto. As construções apresentam as regras que licenciam ou não uma dada expressão linguística de modo a ser bem formada. As manifestações dessas expressões na língua são o que se chama de construto. Em outras palavras, como definem Kay e Fillmore (1999, p. 2-3), a construção consiste em um conjunto de condições que licenciam uma classe de construtos reais de uma língua.

⁴ ... seldom results in a simple concatenation of the meanings those forms might have in isolation.

⁵ Para fins notacionais, os nomes de frames serão apresentados em fonte *Courrier Newe* e os nomes dos Elementos de Frame em *VERSALETE*.

⁶ Para uma análise detalhada da construção morfológica {X-eiro} vide Botelho (2009).

Nas próximas seções, serão apresentadas três das diversas linhas da Gramática das Construções. De tal modo, iniciaremos pela Gramática das Construções de Berkeley, passaremos pela Gramática Cognitivista das Construções e finalizaremos pela Gramática das Construções Corporificada. O motivo pelo qual elegemos essas três vertentes ficará claro quando apresentarmos nossa proposta para a modelagem linguístico-computacional de relações entre construções e entre construções e frames na FrameNet Brasil, uma vez que tal modelagem se dá a partir de reflexões teóricas levadas adiante por essas mesmas vertentes.

1.1 Gramática das Construções de Berkeley

A Gramática das Construções de Berkeley (BCG – *Berkeley Construction Grammar*) tem por objetivo explicar generalizações na língua sem que haja redundância (GOLDBERG, 2006, p. 214). Nesta versão, a gramática de uma língua é compreendida como “o conjunto de suas **construções gramaticais**, as regras que unificam informação formal e semântica em vários tipos de objetos linguísticos” (FILLMORE, 2013, p.112, grifo original).⁷

Kay e Fillmore argumentam que uma das vantagens de uma abordagem construcional unificada é a capacidade com que ela demonstra a interação de construções relativamente idiomáticas com as construções mais familiares para licenciar as sentenças da língua (1999, p.7) – o que os autores demonstram por meio da análise da construção *What's X doing Y?* .

Todas as construções sintagmáticas e muitas das lexicais apresentam uma parte interna e outra externa, conforme explica Fillmore (1999, p.114). A parte externa informa de que tipo são as instâncias licenciadas pela construção, revelando o papel que o sintagma por inteiro pode desempenhar no restante da gramática. A parte interna, por sua vez, descreve os tipos de objetos que instâncias da construção contêm, podendo identificar os tipos de constituintes que podem participar da construção. Assim, a descrição de uma construção de *Sintagma_Nominal_com_Complemento* envolveria tanto informações sobre as propriedades do SN como um todo, quanto informações sobre a necessidade

⁷ ... the set of its **grammatical constructions**, the rules that unite formal and semantic information into various kinds of linguistic objects.

de haver um *NÚCLEO_NOMINAL* que toma um *S_{INTAGMA}_P_{REPOSICIONAL}* como seu argumento.⁸

Nesta proposta, as relações entre elementos linguísticos são formalizadas, a fim de explicar tanto a constituição interna dos construtos, quanto suas propriedades gerais externas. Para tanto, caixas de anotação são utilizadas para organizar as informações contidas em uma dada construção, a qual recebe uma descrição de suas características fonológicas, prosódicas, morfológicas, sintáticas, semânticas, pragmáticas e discursivas. As informações são apresentadas em uma estrutura de traços, na forma de uma matriz de valores de atributos (*AVM – attribute-value matrix*).

Cada atributo pode ter no máximo um valor e qualquer par de AVMs pode ser combinado para licenciar uma dada construção – desde que não haja conflito de valor em nenhum atributo, exceto em casos de coerção de um valor pelo outro. Ao unificarem-se duas AVMs, uma nova AVM é formada contendo os atributos e valores das duas AVMs originais (GOLDBERG, 2006, p.216).

O conceito de unificação é central para a BCG e é representado através da coindexação de elementos presentes nas AVMs. Em linhas gerais, a unificação é o recurso utilizado tanto para se definirem restrições de preenchimento de *slots* da construção, na medida em que apenas constituintes com valores não conflitantes para o mesmo atributo podem coexistir em uma construção (salvo os casos em que a própria construção defina o contrário); quanto para controlar os mapeamentos de valência sintática e semântica da construção, incluindo-se sua composição de significado.

A fim de exemplificar os procedimentos de uma análise construcional da BCG, será descrito, a seguir, um trecho de uma sentença representante da construção de *Dativo_com_Infinitivo_Suficiência* (*DCI_Suficiência*) proposta por Torrent (2015). Tal construção evoca o frame de *Suficiência*, de modo que um ITEM atua como um recurso que possibilita a um beneficiário/agente participar de uma *SITUAÇÃO_HABILITADA*. A maneira como esse ITEM se torna disponível ao beneficiário/agente é determinada pela semântica do verbo finito. Diferentes verbos podem aparecer nessa posição; há

⁸ De modo análogo aos frames, os nomes das construções serão representados em fonte *Courrier New Itálico*, e os nomes de seus elementos constituintes em *VERSALETE ITÁLICO*.

desde verbos de transferência, como “dar”, “colocar” e “passar”, até verbos de atividade mental, como “conhecer”, “saber” e “ler” (vide TORRENT, 2015 para uma descrição completa da construção).

Em relação à sua contraparte sintática, a construção de *Dativo_com_Infinitivo* apresenta o seguinte esquema:

[SN¹ V SN² para (SN³) V_{INF}]

Na construção, o SN³ pode ou não constar, por isso sua representação é feita entre parênteses – o verbo infinitivo terá sempre um sujeito, mas esse poderá ser explícito ou inferido pelo contexto.

Considere-se a sentença (1), exemplar da construção de *DCI_Suficiência*:

(1) Ela deu dinheiro para mim viajar
 SN¹ V SN² para SN³ V_{INF}

Em (1), observa-se que “dinheiro”, um *SINTAGMA_NOMINAL_NÚCLEO* corresponde ao ITEM, e “para mim viajar”, uma *SENTENÇA_INFINITIVA_PARA*, é a *SITUAÇÃO_HABILITADA*, isto é, o item habilita a situação presente na sentença infinitiva:

[SN_Núcleo] → [ITEM]
 [Sent_para_Inf] → [SITUAÇÃO_HABILITADA]

A seguir, será realizada uma análise do trecho “dinheiro para mim viajar” da sentença (1). A descrição da sentença ocorre por meio da inclusão de AVMS em caixas para cada elemento a ser analisado, estando estas caixas em ordem linear conforme aparecem os elementos da sentença. Entretanto, iniciaremos pela descrição de “viajar”, último elemento da sentença, representado na Figura 1.

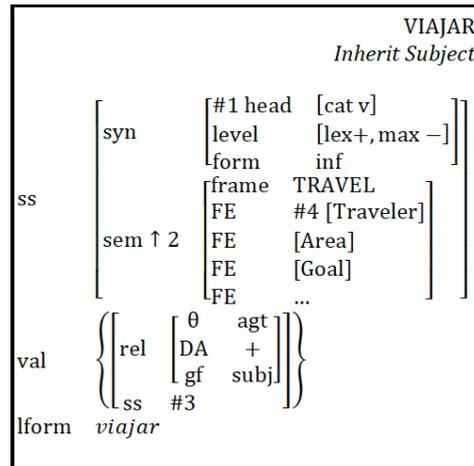


Figura 1: Representação de “viajar”

Nessa representação, observa-se, primeiramente, dentro da especificação sintático-semântica (ss) de VIAJAR a referência “syn” à contraparte sintática. O termo é assinalado como núcleo (“head”) do sintagma, indexado com o número 1 (#1) e de categoria verbal (“cat v”). Em “level” há duas marcações: o fato de se tratar de uma representação lexical (“lex+”) e, também, de não ser uma projeção máxima de “viajar”. Quanto à forma, tem-se um infinitivo (“inf”).

Em relação aos aspectos semânticos, “viajar” remete ao frame de Viagem (“Travel”), no qual há, pelo menos, três Elementos de Frame, a saber, VIAJANTE (“TRAVELER”), indexado com o número 4 (#4), ÁREA (“AREA”) e DESTINO (“GOAL”). Observe que o traço semântico é indexado com o número 2 e uma seta para cima – a qual será explicada adiante. Em relação à valência de “viajar”, esse verbo requer um argumento cujo papel semântico se alinhe a um agente (θ agt) e cuja função gramatical seja sujeito (gf subj), o que faz dele um argumento perfilado, destacado, (“distinguished argument”) para esse verbo (DA +). No que tange à especificação sintático-semântica do elemento que preenche esse slot de valência, temos a indicação de que tal especificação deve ser buscada em outro ponto do diagrama, uma vez que é indexada com o sinal de unificação #3. Aponta-se, ainda, que a forma de manifestação lexical desse construto (“lform”) é “viajar” e, por fim, que, por ser um verbo, “viajar” herda propriedades genéricas da construção de *Sujeito*, que define informações como concordância, por exemplo.

A Figura 2 retrata a sentença infinitiva (*Sinf*) “mim viajar”.

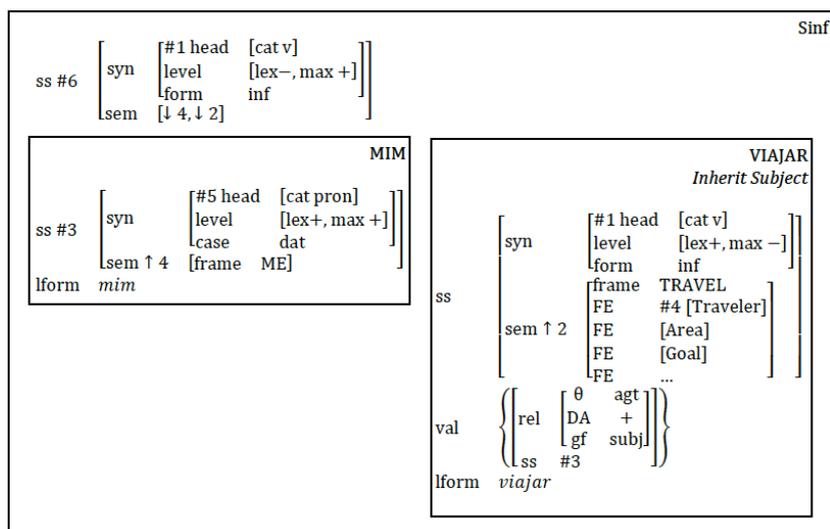


Figura 2: Representação de “mim viajar”

Na Figura 2, observam-se duas caixas menores contidas em uma maior: a caixa com a descrição de VIAJAR, já comentada, e a caixa com a descrição de MIM. Nessa, a especificação *ss* vem unificada àquela prevista pela valência de VIAJAR (#3), indicando que MIM será o sujeito agente do verbo em questão. No que tange à sintaxe, o termo “mim” também é assinalado como núcleo e indexado com #5. Trata-se de um pronome (“*cat pron*”) do caso dativo (“*dat*”) e de uma representação lexical em sua projeção máxima (“*max +*”). Sua contraparte semântica é coindexada ao número 4, uma vez que “mim” é o elemento que cumpre o papel semântico do $V_{IAJANTE}$, assinalado na caixa de VIAJAR, o que exemplifica novamente o processo de unificação.

No topo, à esquerda, vê-se a descrição da sentença infinitiva por completo. O núcleo é coindexado em #1, visto que se trata do elemento “viajar”, de categoria verbal. O sintagma, como um todo, não se constitui em uma representação lexical (“*lex -*”), e essa é sua projeção máxima. O traço semântico apresenta os números 4 e 2 com uma seta, em cada, apontando para baixo; isso indica que a semântica do sintagma corresponde à semântica de cada um dos termos presentes nas caixas – 4 é coindexado à semântica de “mim”, e 2 à semântica de “viajar”.

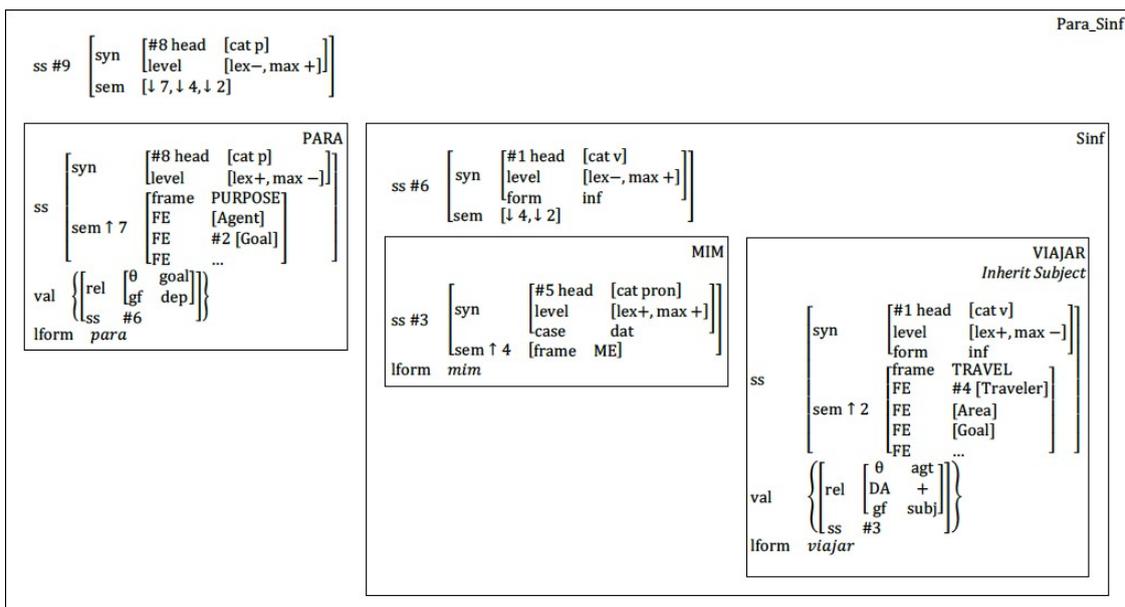


Figura 3: Representação de “para mim viajar”

À Figura 3, acrescenta-se a caixa com o termo “para”, caracterizado como núcleo, da categoria preposição (“cat p”). O elemento é uma representação lexical que não está em sua projeção máxima. O frame evocado é o de Finalidade (“Purpose”), cujos papéis são AGENTE e OBJETIVO, entre outros possíveis. Quanto à valência, é requerido um elemento para o papel θ de OBJETIVO com função gramatical de dependente (“Dep”) e cuja especificação ss unifica-se com aquela da S_{INF} .

No topo, à esquerda, observa-se que o núcleo da sentença infinitiva regida por “para” ($P_{ARA_S_{INF}}$) é atribuído a “para”, pela coindexação #8. O sintagma “para mim viajar” não se constitui em uma representação lexical, e está em sua projeção máxima. Os números 7, 4 e 2 com setas direcionadas para baixo indicam que a semântica do todo é uma composição daquela de cada um dos elementos contidos nas caixas; 7 é coindexado ao frame de Finalidade, e, conforme foi exposto, 4 é coindexado ao frame Mim e 2 a Viagem.

Na Figura 4, PARA primeira caixa, acolhendo outras duas é incluída à esquerda. Na menor delas, aparece o termo “dinheiro”, núcleo, de categoria nominal (“cat n”), de representação lexical, mas não em sua projeção máxima. O frame evocado é o de Dinheiro (“Money”). A próxima caixa a englobar tal termo o enquadra como um sintagma nominal e, por isso, não se trata mais de uma representação lexical, ao passo que a projeção passa a ser máxima.

Na caixa maior, enquadrada como *SINTAGMA_NOMINAL_NÚCLEO* (“*HEAD_NP*”), não se trata mais de uma projeção máxima, uma vez que passa a ter requisitos de valência, e o frame apontado é o de Suficiência (“*Sufficiency*”), o qual possui os EFs ITEM (“*ITEM*”) e SITUAÇÃO_HABILITADA (“*ENABLED_SITUATION*”). Nesse sintagma está contido o item suficiente (dinheiro) – coindexado a #12, o EF “*MONEY*” – para habilitar a situação presente na sentença infinitiva regida por para (“para mim viajar”) – coindexada a #9, correspondente ao *PARA_SINF*. Quanto à valência, é requisitado um elemento de papel θ Objetivo (“*Goal*”) com função gramatical de Dependente (“*Dep*”), também unificado ao *PARA_SINF* (ss #9).

No topo da Figura 4, à esquerda, tem-se que o núcleo de todo o construto se refere ao termo “*dinheiro*”, dada a unificação em #11. Toda a estrutura analisada, de um *Dativo_com_Infinitivo_Suficiência* (“dinheiro para mim viajar”) está em sua projeção máxima e não se trata de uma representação lexical. A contraparte semântica apresenta os números 10, 13 e 7, com setas apontando para baixo, o que aponta que a semântica do *Dativo_com_Infinitivo_Suficiência* alude aos frames de Dinheiro (#10), Suficiência (#13) e Finalidade (#7).

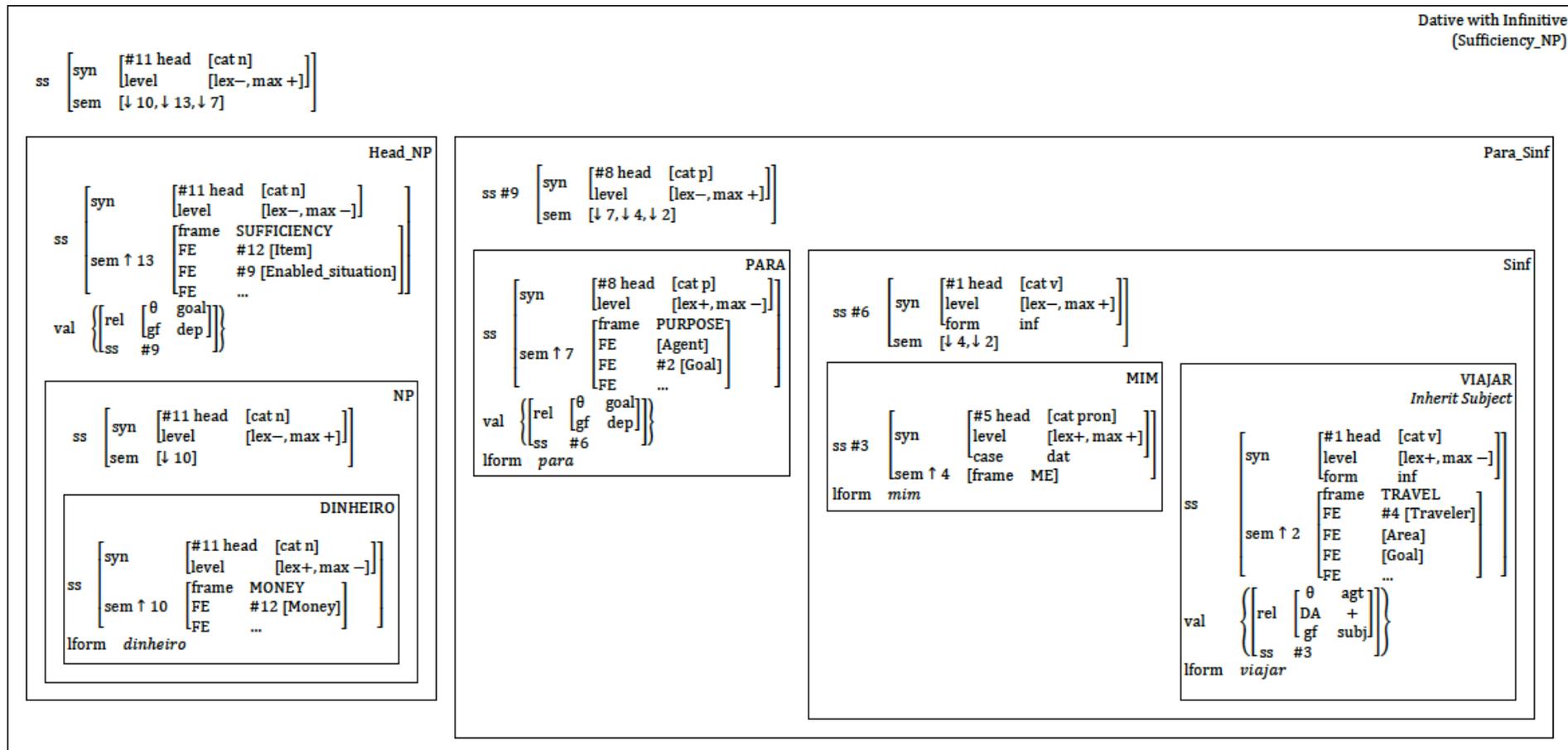


Figura 4: Representação de “dinheiro para mim viajar”

Embora um dos traços previstos para descrição, neste modelo de gramática, seja o semântico, no qual pode-se assinalar o frame a ser evocado pela construção em questão, para Kay e Fillmore (1999), nem toda construção evoca frame, isto é, há aquelas em que se lida apenas com uma construção sintática. Exemplo disso é a construção de *Inversão_de_Sujeito_e_Auxiliar* na língua inglesa, a construção *SAI* (*Subject-Auxiliary_Inversion*).

Fillmore (1999) descreve a construção *SAI* e apresenta suas diferentes instanciações, como se pode observar pelas sentenças de (2) a (7):

- (2) Did you?
- (3) Has anybody seen them?
- (4) Was she there this morning?
- (5) May she live forever!
- (6) Boy, was I stupid!
- (7) Aren't I the clever one!⁹

Essas sentenças têm em comum a inversão do sujeito e do auxiliar, mas, em termos do significado, são divergentes. O autor ressalta que existem algumas funções pragmáticas genéricas em comum em tais frases, como, por exemplo, a de apelação. Entretanto, não é possível saber se essas funções são provenientes do contexto, de propriedades mais genéricas, ou fornecidas pelos próprios itens lexicais. De todo modo, uma vez que frames são cenas das quais fazem parte instrumentos e participantes, não é possível realizar uma definição acerca do frame evocado pela *SAI*. Em outros termos, essa construção não evoca um frame.

Há um desacordo quanto à questão da evocação de frames pelas construções. Enquanto na BCG assume-se que nem toda construção é evocadora de um frame, na Gramática Cognitivista das Construções, por exemplo, afirma-se o oposto, isto é, que toda construção é, efetivamente, um par que une uma forma a um significado, ao frame evocado. Na seção seguinte, esse e outros assuntos serão discutidos a fim de que seja apresentada outra perspectiva de tratamento construcional.

⁹ Sentenças apresentadas por Fillmore (1999).

1.2 Gramática Cognitivista das Construções

Em seu primeiro livro, fruto de sua tese de doutorado, Goldberg (1995, p.1) estabelece uma definição para o termo “construção”, qual seja, um componente que carrega sentido, independentemente das palavras presentes na sentença. Conforme a autora assinala, boa parte do sentido provém de itens lexicais, mas uma abordagem gramatical que se baseie neles, exclusivamente, é falha. Assim, é preciso reconhecer as construções à parte dos itens lexicais que as instanciam.

A fim de confirmar tal afirmação, Goldberg apresenta pares de sentenças cujos itens lexicais são os mesmos, mas a estrutura sintática varia (GOLDBERG, 1995, p.2). Observe um desses pares em (8) e (9).

(8) I brought Pat a glass of water.

(9) I brought a glass of water to Pat.

O enunciado (8) é um caso de sentença *Ditransitiva*, e o (9) sua paráfrase *Dativa* com “to”. Em tais exemplos, a mudança na sintaxe, isto é, a mudança na construção utilizada implica a alteração de sentido. Por exemplo, a sentença *Ditransitiva* exige que seu argumento alvo seja animado, ao passo que o mesmo não corresponde na *Dativa*, conforme (10-11) (GOLDBERG, 1995, p.2):

(10) *I brought the table a glass of water.

(11) I brought a glass of water to the table. (PARTEE, 1965, p. 60 *apud* GOLDBERG, 1995, p.2)

Além desse, diversos casos com mudanças de significado provenientes de mudanças na estrutura argumental foram observados, como o das sentenças (12-13), estudadas por Fillmore (1968, p.49):

(12) Bees are swarming in the garden.

(13) The garden is swarming with bees.

Na sentença (13), subentende-se que há muitas abelhas em todo o jardim, enquanto em (12) as abelhas poderiam estar presentes em apenas uma parte do jardim.

Abordagens por regras lexicais, como sinalizado por Goldberg (1995, p.9), aproximam-se da abordagem construcional da autora, na medida em que enfatizam as diferenças semânticas geradas a partir de diferenças na configuração do complemento. Pinker (1989 *apud* GOLDBERG, 1995), ao postular regras que operam na estrutura semântica, em oposição a regras puramente sintáticas, também apreende que mudanças na estrutura do complemento são essencialmente semânticas. Contudo, as duas abordagens diferem da abordagem construcional de Goldberg, conforme essa avulta a relação entre verbo e construção – o que discutiremos posteriormente.

Contrariamente à postulação de Fillmore (1999), Goldberg afirma que toda construção evoca um frame. A pesquisadora defende que “motivações funcionais sincrônicas se escondem atrás de fatos sintáticos aparentemente brutos” (GOLDBERG, 2006, p. 167), alegando que não se reconhece uma generalização funcional acerca da construção *SAI* devido a uma (falsa) suposição de que deve haver um único traço ou um grupo de traços comuns à categoria.

Diversas análises levantadas por Goldberg (2006) – tais como as de Lakoff & Brugman (1987), de Lambrecht, (1994) e de Michaelis & Lambrecht (1996) – observaram que a construção *SAI* é um desvio da sentença prototípica, sendo esta última declarativa, uma asserção positiva com estrutura informacional com foco no predicado e independente de outra oração (GOLDBERG, 2006, p. 168). Entretanto, nenhuma construção *SAI* possui todos os atributos das sentenças não-prototípicas, o que se concilia com as normas de categorização: “um subconjunto de atributos que derivam do caso prototípico é instanciado em cada extensão convencionalizada” (GOLDBERG, 2006, p. 169).¹⁰ Segundo Goldberg (2006, p. 176), a maioria dessas construções apresenta ao menos duas características das sentenças não-prototípicas.

¹⁰ ... a subset of attributes that hold of the prototypical case is instantiated in each conventionalized extension.

Assim, o protótipo a partir do qual as construções *SAI* manifestam-se é uma sentença não-prototípica, a qual é não-positiva, não tem foco em predicado, não é assertiva, é dependente e não-declarativa. Contudo, tal sentença é uma generalização que não pode ser diretamente instanciada (GOLDBERG, 2006, p. 176).

Conforme se mostrará no restante desta tese, o Constructicon da FrameNet Brasil adota a perspectiva Fillmoreana para a relação entre construções e frames, mapeando a contraparte semântica das construções apenas quando ela pode, de fato, ser representada em termos de uma cena composta por participantes e os instrumentos que utilizam (conforme a definição de frame proposta por FILLMORE, 1982).

Goldberg (1995) aponta que muitas construções são inter-relacionadas, isto é, podem ser associadas a uma família de sentidos relacionados, apesar de distintos (GOLDBERG, 1995, p.4). Para a linguista, uma compilação de construções configura-se como uma rede altamente estruturada de informação relacionada. Argumenta que construções compõem uma rede de relações de herança, por meio das quais é possível depreender generalizações e exceções.

Para melhor compreensão das relações entre construções, faz-se necessária a descrição dos princípios psicológicos de organização da língua:

I – *O Princípio da Motivação Maximizada*: Se uma construção A é relacionada sintaticamente à construção B, então o sistema da construção A é *motivado* na medida em que é relacionado semanticamente à construção B (...). Tal motivação é maximizada.

II – *O Princípio da Não Sinonímia*: Se duas construções são distintas sintaticamente, elas devem ser distintas semântica ou pragmaticamente (...). Aspectos pragmáticos de construções envolvem elementos de estrutura de informação, incluindo tópico e foco e, adicionalmente, aspectos estilísticos da construção tais como registro. (...)

III – *O Princípio da Expressividade Máxima*: O inventário de construções é maximizado para fins comunicativos.

IV – *O Princípio da Economia Máxima*: O número de construções distintas é minimizado tanto quanto possível, tendo em conta o Princípio III (...). (GOLDBERG, 1995, p.67-68, grifo original).¹¹

¹¹ I. *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.

II. *The Principle of No Synonymy*: if two constructions are syntactically distinct, they must be semantically or pragmatically distinct (...). Pragmatic aspects of constructions involve particulars of information structure, including topic and focus, and additionally stylistic aspects of the construction such as register. (...)

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

IV. *The Principle of Maximized Economy*: the number of distinct constructions is minimized as much as possible, given Principle III (...).

Para melhor compreensão dos princípios, Goldberg esclarece, à luz de analogias realizadas por Haiman¹² (1985 *apud* GOLDBERG, 1995, p.68), que diferenças na forma ocasionam diferenças no significado – o que aponta para o Princípio da Não-Sinonímia. Em contrapartida, diferenças no significado deveriam levar a diferenças na forma – o que é capturado pelo Princípio da Expressividade Máxima. As exceções a esse preceito relacionam-se a casos de polissemia e homonímia, contudo são atribuídas a uma necessidade de simplificação – Princípio da Economia Máxima. Assim, enquanto o último princípio restringe o número de construções possíveis, o Princípio da Expressividade Máxima cria a disposição para mais formas distintas, isto é, os dois princípios restringem-se mutuamente (GOLDBERG, 1995, p.69).

Goldberg (1995, p. 69) propõe que, enquanto os Princípios II, III e IV são amplamente assumidos, o Princípio da Motivação Maximizada não compartilha de tal assentimento e, por essa razão, a discussão sobre ele é ampliada. Argumenta-se que a motivação é uma propriedade fundamental para o funcionamento das línguas, uma vez que permite generalizações e simplificações, condições fundamentais dada a memória finita do ser humano. Goldberg (2006, p.217) assinala, ainda, que o conceito de motivação é utilizado a fim de explicar a existência de um dado pareamento de forma e significado em uma dada língua. Segundo a linguista, esse conceito é fundamental para a Gramática Cognitivista das Construções, a qual busca fornecer tal componente para cada construção postulada.

A linguista explica (Goldberg, 1995, p.69) que o papel da motivação jaz entre a previsibilidade e a arbitrariedade. Ilustra essa disposição a palavra *dix-neuf*, do francês, a qual remete ao número “19”. Embora *dix* refira-se ao número “dez”, e *neuf*, ao número “nove”, não se pode prever que tal número tomaria essa exata forma na língua francesa – visto que poderia ser composta, por exemplo, a forma *neuf-dix*. Em contrapartida, não se pode alegar que o termo seja arbitrário – já que *dix* (“dez”) somado a *neuf* (“nove”) resulta em *dix-neuf* (“dezenove”).

¹²Haiman compara a forma de uma língua a um diagrama, como, por exemplo, um mapa. Enquanto esse representa a geografia, a língua representa a nossa construção da realidade.

O conceito de herança tem sido empregado para capturar generalizações linguísticas. Assim, as relações de motivação são identificadas por meio de links de herança assimétrica entre construções relacionadas semanticamente e sintaticamente, isto é, a construção A motiva a construção B se B herda de A.

A seguir, observe a representação de uma relação de herança entre duas construções, C1 e C2:

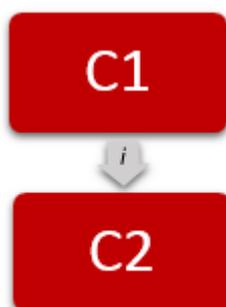


Figura 5: Relação de herança entre construções

Na Figura 5, C2 herda de C1, e C1 domina C2. Pode-se sinalizar, também, que C1 motiva C2. A letra “i”, na seta, representa o link de herança entre as construções.

Goldberg (1995, p.73-74) assume o modelo de herança múltipla, no qual uma construção, na hierarquia, pode herdar de mais de uma construção dominante. Além disso, adota-se o modo normal de herança, o qual permite exceções e generalizações parciais. Nesta gramática, as relações de herança são compreendidas, também, sob a ótica da cópia real de informação, pela qual construções dominadas contêm toda a informação que as construções dominantes apresentam. Assim, explica a pesquisadora, cada construção é especificada por completo, mas é redundante na medida em que a informação é herdada de construções dominantes.

As relações entre construções são pautadas por diferentes links de herança, os quais demonstram que toda informação não conflitante entre duas construções relacionadas é compartilhada (GOLDBERG, 1995, p.74-75). Tais links serão explorados adiante, na seção 3.1.

A seguir, será apresentada uma terceira abordagem construcional para a gramática, a qual acrescenta novos e importantes *insights* para o desenvolvimento da presente pesquisa.

1.3 Gramática das Construções Corporificada

Bergen e Chang (2013, p.169) esclarecem que a Gramática das Construções Corporificada (ECG – *Embodied Construction Grammar*) busca explicar os mecanismos e representações utilizados pelos seres humanos na aprendizagem, produção e compreensão de enunciados. Tal concepção demonstra que essa perspectiva gramatical se aproxima mais da ciência cognitiva que propriamente da linguística analítica.

Assim sendo, a pergunta norteadora da ECG pode ser sintetizada da seguinte forma: “Quais mecanismos cognitivos e neurais são mobilizados pelas pessoas quando fazem uso da linguagem humana?” (BERGEN & CHANG, 2013, p.169)¹³, ou seja, essa gramática investiga a base neurológica da linguagem. Embora a ECG siga o ciclo científico convencional de investigação – partindo de observações, para, então, levantar hipóteses e, a partir delas, construir modelos e avaliá-los –, a ênfase não é atribuída à aparência das construções, mas ao modo como elas são usadas, uma vez que construções não são apenas objetos descritivos.

Como esse modelo também é voltado para formalizações computacionais, assim como a BCG, é considerado mais detalhado que a Gramática Cognitivista das Construções, apesar de compartilhar com esta o viés psicolinguístico.

Para se saber como o significado se constitui, a hipótese lançada é a de que o indivíduo compreende os enunciados por meio de simulações corporificadas. Nas últimas décadas, estudos voltados para o processamento (acesso e construção) do significado têm apontado para a conclusão de que quando as pessoas processam palavras ou sentenças, simulações corporificadas, visuais ou motoras, das cenas mencionadas são ativadas em suas mentes (BERGEN, 2012). Assim, o processo de compreensão das sentenças parece envolver a ativação de representações internas dos eventos perceptíveis ou das ações executáveis.

¹³ What cognitive and neural mechanisms do people engage while using human language?

Na ECG, palavras e construções são caminhos que conectam o conhecimento sobre a forma ao conhecimento sobre o significado. Tais caminhos são mediados por categorias, as quais se constituem em esquemas e construções. Os esquemas são representações que generalizam sobre instâncias; por exemplo, ações motoras muito diferentes são esquematizadas como “entrar”, isto é, há um esquema executor dessas ações. As construções, por sua vez, são estruturas inteiras que conectam os esquemas.

Bergen e Chang (2013, p.174) elucidam que para uma descrição do significado fundamentada em simulação corporificada é necessário recorrer à Semântica da Simulação (NARAYANAN, 1997; FELDMAN, 2006), a qual propõe que os mesmos sistemas dinâmicos de controle usados para executar ações ou perceber eventos são usados para simular ações e eventos.

Dodge (2010, p. 43) também apresenta os conceitos centrais empregados pela ECG, o conceito de construções e o de esquemas, porém, com vias a discutir sua representação computacional. A autora define as construções como pareamentos de restrições de forma e de restrições de significado e argumenta que a ECG difere de outras perspectivas construcionais na medida em que as restrições de significado são definidas em termos de esquemas semânticos corporificados.

Os esquemas são utilizados para representar diversas estruturas conceptuais, tais como os esquemas imagéticos e os frames – que são, por sua vez, utilizados estritamente pela BCG e menos pontualmente pela Gramática Cognitivista das Construções – e pertencem a uma grade de esquemas inter-relacionados. Dodge (2010, p.44-47) apresenta os esquemas através de exemplos, como mostra a Figura 6:

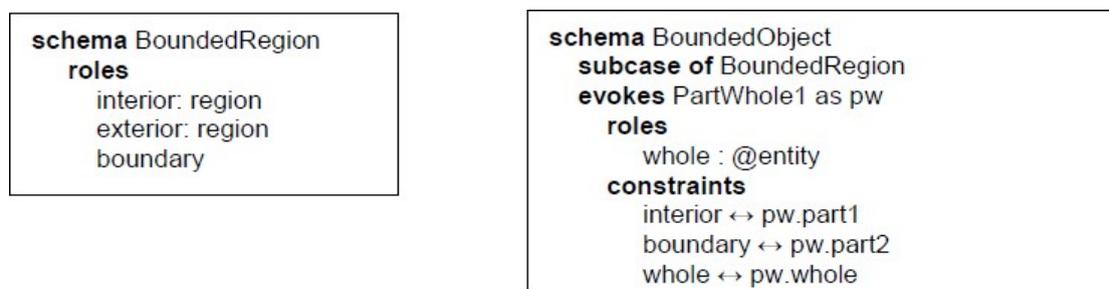


Figura 6: Os esquemas BoundedRegion e BoundedObject (DODGE, 2010, p. 46)

No primeiro quadro, observa-se o esquema BoundedRegion, o qual remete a uma região delimitada, cujos papéis são: interior (região do espaço delimitada por uma fronteira), exterior (região fora da fronteira) e fronteira (“*boundary*”, a qual delimita sua região interior). Os dois primeiros papéis possuem restrições, já que ambos são do tipo “região” (*region*). Segundo a pesquisadora explica, tal esquema representa os elementos estruturais chave de uma conceptualização esquemática particular de alguma extensão do espaço (DODGE, 2010, p. 45).

O segundo quadro, BoundedObject, é um subcaso do esquema BoundedRegion, uma vez que contém toda a estrutura do esquema "pai", mas também especifica estrutura adicional. Este novo esquema descreve objetos delimitados, tais como contêineres, os quais possuem duas partes: uma fronteira e uma região interior – papéis herdados de BoundedRegion (DODGE, 2010, p. 45).

Além dos papéis que são herdados, BoundedObject ainda apresenta outro, o todo (“*whole*”), e evoca o esquema PartWhole, o qual recebe a notação “pw”. O termo “pw.part1” remete ao papel “part 1” desse último esquema; assim, a restrição “interior ↔ pw.part1” aponta que o papel “interior” de BoundedObject é ligado ao papel “part1” de PartWhole, formando um papel mais complexo. Desse modo, “interior” e “fronteira” são partes de um todo maior (DODGE, 2010, p. 45).

Dodge (2010, p. 47) explica que se assume a existência da estrutura conceptual representada pelos esquemas acima descritos, independentemente de qualquer construção linguística específica. Os significados das construções podem ser representados pelos esquemas em várias combinações, usando-se, por exemplo, mais de um esquema. Do mesmo modo, o mesmo esquema pode representar o significado de mais de uma construção.

Na Figura 7, será demonstrada a representação de uma construção lexical feita por Dodge segundo o formalismo da ECG, a fim de evidenciar os procedimentos analíticos dessa teoria gramatical:

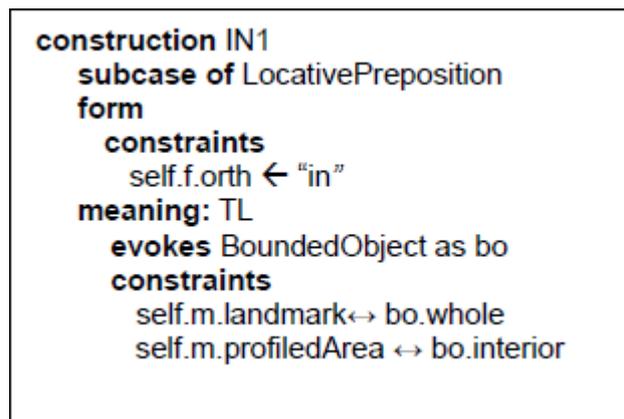


Figura 7: A construção "in" (DODGE, 2010, p. 49)

No topo do quadro, observa-se o nome da construção: "IN1", a qual é nomeada de acordo com a sua forma ortográfica ("*in*"), associada a um número ("1"), visto que a mesma forma pode ser vinculada a mais de um significado. Em seguida, há a informação de que se trata de um subcaso da construção LocativePreposition (preposição locativa), a qual generaliza sobre diversas preposições espaciais, capturando semelhanças no significado e no uso de várias construções (DODGE, 2010, p. 48).

Em seguida, iniciam-se as informações acerca da forma e do significado da referida construção. Quanto à forma, a seguinte restrição é sinalizada: "self.f.orth ← 'in'". Dodge (2010, p. 48) explica que o termo "self" indica que a construção será considerada como um todo; "f" sinaliza que se contemplam mais especificamente os elementos da forma da construção; e "orth" é usado para referir-se à forma ortográfica; a seta especifica uma restrição de valor, o qual é, neste caso, a sequência ortográfica "in".

No tangente ao significado, a construção como um todo é identificada como esquema TL (*trajectorlandmark* – trajetor ponto de referência) como uma restrição de tipo. Dodge (2010, p. 48) esclarece que esse esquema é acionado uma vez que "*in*" especifica uma localidade, a qual é definida em relação a um dado marco. Tal esquema, conforme será mostrado na Figura 8, apresenta três papéis: trajetor, ponto de referência e área perfilada:



Figura 8: O esquema TL (DODGE, 2010, p. 46)

Depois de assinalar a identificação da construção “*in*” com o esquema TL, nota-se que essa evoca o esquema BoundedObject (Figura 6), pois o ponto de referência é conceptualizado como um objeto delimitado, e a localização relevante é a região interior desse objeto (DODGE, 2010, p. 48).

Nas “restrições”, são demarcadas as relações entre os dois esquemas mencionados. Primeiramente, o papel “ponto de referência” do esquema TL é vinculado ao papel “todo” de BoundedObject, a fim de evidenciar que o ponto de referência é conceptualizado como um objeto delimitado. Em seguida, o papel “área perfilada” do TL é vinculado ao papel “interior” do BoundedObject, já que a localização relevante é o interior do marco.

Como se pode notar, assim como nas demais abordagens da Gramática das Construções aqui apresentadas, a ECG também trabalha com construtos teóricos específicos para lidar com as relações entre construções e entre estas e os frames para que apontam. No capítulo 3 voltaremos a discutir essas questões em maior profundidade. Damos sequência à apresentação dos conceitos fundamentais necessários a esta tese no capítulo seguinte, em que situamos a pesquisa desenvolvida no domínio da Linguística Computacional.

2 LINGÜÍSTICA COMPUTACIONAL: FUNDAMENTOS E IMPLEMENTAÇÕES COGNITIVAMENTE INSPIRADAS

2.1 Fundamentos da Linguística Computacional

Os trabalhos em Linguística Computacional começaram a se desenvolver na década de 1950, com foco principal em tradução automática. Na década seguinte, é possível afirmar, conforme assinala Spärck-Jones (2007, p.439), que os linguistas chomskianos e os computacionais compartilharam uma crença geral na importância de um aparato formal que pareasse cadeias de palavra com descrições de estrutura gramatical.

Posteriormente, surgiu uma divergência entre as duas áreas, de modo que, sem muita contribuição da linguística tradicional, a vertente computacional prosseguiu com suas investigações na década de 1970, apesar da perda de financiamento para a tradução automática (SPÄRCK-JONES, 2007, p.439).

O afastamento entre a Linguística e a sua modalidade computacional permaneceu na década de 1980 e se tornou mais manifesto na de 1990. Essa modalidade desenvolveu pesquisas com base em *corpus*, as quais incorporavam ideias probabilísticas, realizando a sua própria contribuição para os modelos de linguagem. Assim, ao menos em sua abordagem *mainstream*, não é raro que os trabalhos atuais em Linguística Computacional se baseiem em modelos puramente matemáticos, tais como os aplicados à tradução estatística (KOEHN, 2009) e à representação vetorial de itens linguísticos (GOLDBERG & LEVY, 2014).

Entretanto, há modelos de representação computacional das línguas naturais que ainda consideram alguma ancoragem linguística. Tais modelos baseiam-se na premissa de que a representação de aspectos da estrutura linguística – ou cognitiva – podem contribuir para as tarefas de Compreensão de Língua Natural, mesmo as que usam métodos estatísticos.

Ibáñez e Masegosa (2014) propõem os requisitos para um modelo computacional inspirado nos trabalhos do campo da Linguística Cognitiva sobre metáforas, metonímias, demais processos semânticos e sua manifestação linguística. O modelo proposto, chamado Modelo Léxico-Construcional (MLC), adota uma concepção construcionista de língua baseada no uso, isto é,

suas análises seriam fundamentadas em ocorrências de expressões linguísticas dentro de seu contexto espontâneo de produção (2014, p.30). Nesse sentido, tal proposta aproxima-se de implementações já correntes, tais como o ECG Analyzer (BRYANT, 2008) e a FrameNet (FILLMORE ET AL., 2003), conforme se demonstrará nas seções 3.3 e 2.3, respectivamente.

São propostos pelo MLC quatro níveis em seu aparato descritivo. No primeiro nível, o da camada da estrutura argumental, se encontraria o modelo lexical, que, conforme Ibáñez e Masegosa explicam (2014, p.31), capturaria estruturas como o frame, o qual é relacionado à estrutura de evento de verbos e de outros predicados. Encontrar-se-ia, ainda, nesse nível, o modelo construcional de estrutura argumental, cuja atividade origina representações de estrutura de evento, enriquecidas semântica ou inferencialmente, as quais estão prontas para interação com representações de outros níveis. Enquanto o modelo lexical seria uma representação proposicional de nível baixo, o segundo modelo seria de nível alto, tal qual as construções *Ditransitivas* e *Resultativas*, por exemplo.

O segundo nível abordaria o significado com base em uma situação. Ibáñez e Masegosa (2014, p.31) assinalam que tal significado pode associar-se convencionalmente a uma forma linguística, ocasionando, assim, construções implicacionais, as quais surgem a partir de implicações de significado derivadas originalmente por meio de mecanismos inferenciais. Exemplo disso são as sentenças da construção *WXDY*¹⁴, como em “O que essa toalha está fazendo em cima da cama?”. O falante solicita uma resposta ao ouvinte, mesmo já estando ciente dessa. A estranheza da situação origina a implicação de que o fato de a toalha estar nesse local é inadequado.

O terceiro e o quarto níveis corresponderiam às dimensões pragmática e do discurso.

A proposta de Ibáñez e Masegosa (2014) não apresenta uma implementação propriamente dita, mas aponta requisitos que devem ser considerados em um modelo computacional que se proponha incluir aspectos da relação entre língua e cognição. Dentro do paradigma da Linguística Computacional, o MLC localiza-se no nível de análise representacional,

¹⁴ Construção *What's X Doing Y?* (KAY & FILLMORE, 1999).

conforme definiremos a seguir. Na proposição das relações objeto desta tese, os pressupostos da modelagem de base cognitiva são respeitados, como se mostrará no capítulo 5, uma vez que tanto construções de nível alto, quanto as de nível baixo têm seu aporte semântico modelado em termos de frames.

2.2 Níveis de Análise

A complexidade da linguagem humana é rapidamente constatada ao se investigar tal objeto; uma série de etapas de análise se evidencia compondo o que chamamos de gramática da língua. Paradoxalmente, logo nos primeiros anos de vida, um indivíduo é capaz de apreender esse conjunto de preceitos e dominá-lo com maestria. Deveria, portanto, ser menos trabalhoso fazer o computador compreender uma língua. Entretanto, esse empreendimento é custoso, uma vez que ainda se sabe pouco sobre como modelar o aparato cognitivo humano utilizado para a aquisição e uso do conhecimento linguístico.

Nesse sentido, os estudos em Processamento de Língua Natural (PLN) podem contribuir para o próprio entendimento das línguas naturais, já que incentivam – e necessitam – (d)a reflexão sobre elas:

Os modelos [de PLN] são necessariamente incompletos [...] Mas, mesmo assim, constituem um referencial claro por meio do qual podemos refletir sobre o que é que fazemos quando compreendemos uma língua natural ou reagimos aos atos de fala nela codificados. (WINOGRAD, 1972 apud DIAS-DA-SILVA ET AL., 2007, p.12)

Os níveis de análise que serão descritos nesta seção revelam as tarefas a serem executadas pelo profissional de PLN. É importante esclarecer que, ao se desenvolver pesquisas nesta área, tem-se por objetivo facilitar a comunicação entre o ser humano e os computadores via códigos linguísticos, por meio da implementação de sistemas computacionais.

Assim, as investigações linguísticas no âmbito do PLN apresentam diversas aplicações, a saber, tradução de texto, revisão ortográfica, análise sintática, entre outras. É por esse motivo que Dias-da-Silva et al. (2007) afirmam que os estudos nessa área podem ser concebidos como uma engenharia do conhecimento humano. Como tal, o processo de formação de um sistema de PLN (SPLN) envolve três etapas de desenvolvimento:

- **Fase Linguística:** construção do corpo de conhecimentos sobre a própria linguagem, dissecando e compreendendo os fenômenos linguísticos necessários para o desenvolvimento do sistema. Nesta fase, a análise dos fenômenos linguísticos é elaborada em termos de modelos e formalismos desenvolvidos no âmbito da teoria linguística.
- **Fase Representacional:** construção conceitual do sistema, envolvendo a seleção e/ou proposição de sistemas formais de representação para os resultados propostos pela fase anterior. Nesta fase, projetam-se as representações linguísticas e extralinguísticas em sistemas formais computacionalmente tratáveis.
- **Fase Implementacional:** codificação das representações elaboradas durante a fase anterior em termos de linguagens de programação e planejamento global do sistema. Nesta fase, além de transformar as representações da fase anterior em programas computacionais, estudam-se as questões referentes à integração conceitual e física dos vários componentes envolvidos, bem como questões referentes ao ambiente computacional em que o sistema será desenvolvido e implementado. (DIAS-DA-SILVA ET AL., 2007, p.13)

Conforme explicam os autores, na Fase Linguística, a primeira delas, trabalha-se com questões referentes ao domínio linguístico, como conceitos, regras e princípios, a fim de explicitar e especificar conhecimentos sobre a língua em questão. A Fase Representacional, por sua vez, é responsável por propor sistemas de representação formal que compreendem, por exemplo, a lógica, redes semânticas e frames. Por fim, a Fase Implementacional cuida da representação, em programas, do produto das fases anteriores e da própria programação do sistema computacional a ser utilizado na tarefa em foco (DIAS-DA-SILVA ET AL., 2007, p.13).

Os autores propõem ainda que as fases supracitadas se desenvolvam de forma sucessiva, progressiva e cíclica. É sugerida a dinâmica representada na Figura 9.

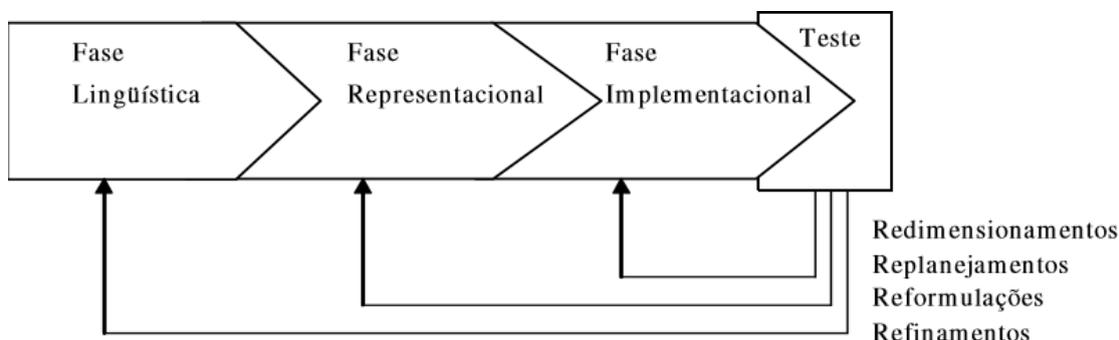


Figura 9: A dinâmica do processo de construção de um SPLN (DIAS-DA-SILVA ET AL., 2007, p.14)

Desse modo, para se completar um ciclo, o produto das duas primeiras fases deve ser implementado na terceira fase para, finalmente, ser testado. As fases, por sua vez, poderão ter seus resultados aprimorados por meio de testes de adequação e desempenho (DIAS-DA-SILVA ET AL., 2007, p.14).

O trabalho desenvolvido nesta tese parte de análises realizadas por outros estudiosos e as utiliza como produto da Fase Linguística. São revisitados os trabalhos de Torrent (2009, 2015) sobre a família de construções *Para Infinitivo* e Sigiliano (2012, 2013) sobre as Construções Aspectuais Inceptivas. Debruça-se, focalmente, sobre a Fase Representacional ao propor especificações para a modelagem de relações entre construções e entre estas e os frames no âmbito da FrameNet Brasil. Por fim, utiliza-se da Fase Implementacional, através da modelagem das relações no banco de dados da FrameNet Brasil.

Dado esse contexto, passamos, a seguir, à apresentação da FrameNet, tanto no que concerne ao Lexicon quanto ao Constructicon, uma vez que será este o *locus* implementacional da pesquisa desenvolvida nesta tese.

2.3 Implementações Computacionais Cognitivamente Inspiradas

2.3.1 Lexica Baseados em Frames

Em 1997, no International Computer Science Institute (Berkeley, EUA), inicia-se oficialmente o projeto FrameNet por iniciativa de Charles Fillmore e Beryl Atkins. A Semântica de Frames passaria a ser empregada como base teórica para tal empreendimento lexicográfico-computacional, possibilitado, naquele momento, pelo avanço das tecnologias de informação.

O projeto de Berkeley buscou, desde o seu início, descrever as propriedades semânticas e distribucionais de palavras do inglês utilizando-se de sentenças extraídas de *corpora* para anotação. Ademais, o recurso foi criado com o objetivo de contribuir para tarefas em Compreensão de Linguagem Natural – ou NLU, do inglês *Natural Language Understanding* (ALLEN, 1995) –, uma vez que foi idealizado com base em uma metodologia que admite o tratamento computacional.

Principal suporte teórico da Berkeley FrameNet, a mencionada hipótese semântica apresenta como elemento central o conceito de frame, o qual emana de postulações teóricas originárias da Inteligência Artificial (MINSKY, 1975) e da Sociologia Interacional (GOFFMAN, 1974). O referido conceito assume como princípio que “todo processo de significação linguística constitui o enquadramento contextualizado (*framing*) de uma situação, que se apresenta estruturada por uma constelação de elementos que a distinguem de outras situações” (SALOMÃO, 2013, p.9).

Assim, observemos como essa teoria é aplicada na FrameNet, no contexto real de uma sentença. Ao se considerar a lexema “estudar”, em (14), é acionado cognitivamente o frame de *Estudar*, cuja definição é apresentada em seguida.

- (14) Ele passou dez anos **estudando** física na Universidade antes de ser reprovado.

Um **Estudante** se matricula e, então, permanece em uma **Instituição** com a finalidade de ser instruído em uma **Matéria**. Ele pode receber instrução de um determinado **Professor** na **Instituição**. O **Local**, o **Tempo** e a **Duração** do estudo também podem ser especificados¹⁵.

Ao item evocador do frame, atribui-se o nome de Unidade Lexical (UL), componente que pode ser definido como o pareamento de um lema a um frame. Em outros termos, uma UL é “uma ‘palavra’ tomada em um de seus sentidos” (FILLMORE ET AL., 2003, p. 297)¹⁶. Logo, nessa exemplificação, a UL “estudando” evoca a cena acima exposta. Assim, cada descrição lexical na FrameNet conecta uma UL ao frame por ela evocado.

A supracitada “constelação de elementos”, responsável por estruturar e distinguir o evento em relação aos demais, vem marcada na definição acima em cores. Cada termo destacado – como ESTUDANTE ou PROFESSOR – corresponde a um Elemento de Frame (EF), isto é, trata-se dos termos constituintes da sentença de acordo com os papéis semânticos que assumem em relação à unidade evocadora do frame. No caso do frame de *Estudar*, os EFs nucleares,

¹⁵ Disponível em <http://www.ufjf.br/framenetbr/>

¹⁶ ... a ‘word’ taken in one of its senses.

isto é, os elementos necessários para a instanciação da cena, contemplam os seguintes papéis:

Estudante: Aquele que recebe instrução de um **Professor** ou **Instituição**.
Instituição: Um estabelecimento educacional, como uma escola ou uma universidade.
Matéria: A área do conhecimento ou habilidade que é ensinada a um **Estudante**.
Professor: Aquele que instrui um **Estudante** em alguma área do conhecimento ou em alguma habilidade¹⁷.

Logo, é possível identificar tais EFs na sentença (14), marcando-os da mesma cor usada em sua descrição acima. A UL, por sua vez, aparece escrita em branco com fundo negro, conforme a Figura 10, extraída da ferramenta de anotação e gestão de banco de dados da FrameNet Brasil, a WebTool.



Figura 10: EFs presentes na sentença (14)

Conforme se observa, um dos quatro EFs anteriormente descritos não consta na sentença, qual seja, o EF PROFESSOR. É importante ressaltar que EFs podem estar em Core Set, isto é, podem atuar em conjunto, de modo que a presença de algum membro do conjunto seja suficiente para satisfazer a valência do predicador (RUPPENHOFER ET AL., 2016, p.25). Esse é o caso dos EFs PROFESSOR e INSTITUIÇÃO: se um for mencionado, não é necessário que haja uma referência ao outro. Como INSTITUIÇÃO consta na sentença, PROFESSOR não precisa aparecer. Por outro lado, foram marcados dois EFs que não haviam sido mencionados anteriormente, a saber, DURAÇÃO e TEMPO, os quais se configuram como EFs não nucleares.

Assim sendo, se alguém realizar a busca por um frame na FrameNet, o encontrará seguido de uma definição em forma de texto, a qual caracteriza a cena ampla evocada. Constarão, ainda, as definições acerca dos participantes do frame, os EFs, conforme se verifica na Figura (11).

¹⁷ Disponível em <http://www.ufjf.br/framenetbr/>

Estudar

Definição
Um **Estudante** matricula e, em seguida, permanece em uma Instituição com a finalidade de ser instruído em alguma Matéria. Ele pode receber instrução de um determinado **Professor** em uma Instituição. O **Lugar**, **Tempo** e a Duração do estudo podem também ser especificados. Ele passou dez anos estudando física na Universidade antes de ser reprovado.

Exemplo(s)

Elementos de Frame Nucleares

Estudante [student] Aquele que recebe instrução de um **Professor** ou Instituição.

Instituição [institution] Um estabelecimento educacional, como uma escola ou uma universidade.

Matéria [subject] A área do conhecimento ou habilidade que é ensinada a um **Estudante**. Ela estuda astrofísica.

Professor [teacher] Aquele que instrui um **Estudante** em alguma área do conhecimento ou em alguma habilidade. Ela estudou com ele por anos.

FE Core set(s):
{Instituição, Professor}

Elementos de Frame Não-Nucleares

Co_participante [co_participant] Co_participante é uma entidade que participa de um forma coordenada em um evento de Estudar.

Descritor [deictive] Um estado que descreve o Estudante durante o processo de estudo.

Duração [duration] O período de tempo em que um processo de estudo é desenvolvido.

Explicação [explanation] A razão pela qual o **Estudante** estuda uma dada Matéria.

Lugar [place] O **Lugar** é o local onde o estudo ocorre. NOTA: Observe se um possível sintagma de **Lugar** é realmente onde o ensino ocorre ou, em vez disso, parte de uma descrição da Instituição.

Maneira [manner] A maneira de estudar que um **Estudante** tem.

Motivação [motivation] A Motivação do **Estudante** para estudar. Ele está estudando para passar no mestrado .

Nível [level] Este EF identifica o Nível de um **Estudante** no seu processo de educação. Um **Estudante** de graduação Um **Estudante** de quarta-série

Tempo [time] O tempo em que o Estudante está estudando.

Figura 11: Frame de *Estudar* (<http://www.framenetbr.ufjf.br>)

Sentenças extraídas de *corpora* diversos são anotadas na FrameNet a fim de se analisar de que modo ULs evocam frames e como EFs são instanciados na prática. Além disso, as propriedades sintáticas presentes na sentença também são anotadas – como será mostrado adiante.

Considerando-se essas definições, a anotação na FrameNet pode ocorrer por meio de duas metodologias: a lexicográfica e a de texto corrido. Na primeira, inicia-se pela escolha do frame a ser trabalhado, descrevendo-se sua estrutura conceptual. Essa descrição deve:

- (i) caracterizar esquematicamente o tipo de entidade ou situação representada pelo frame, (ii) escolher mnemônicos para rotular as entidades ou componentes do frame, e (iii) construir uma lista de trabalho das palavras que parecem pertencer ao frame, sendo que a adesão ao mesmo frame significará que as sentenças que contêm as ULs permitirão análises semânticas comparáveis (FILLMORE ET AL., 2003, p. 297).¹⁸

¹⁸ (1) characterizing schematically the kind of entity or situation represented by the frame, (2) choosing mnemonics for labeling the entities or components of the frame, and (3) constructing a working list of words that appear to belong to the frame, where membership in the same frame will mean that the phrases that contain the LUs will all permit comparable semantic analyses.

Após a realização da terceira tarefa, qual seja, a de definir a lista de ULs alusivas ao frame em questão, inicia-se uma análise de cada uma delas. Primeiramente, são avaliados os diversos usos de uma UL encontrados nos resultados de busca no *corpus*, separando-se os casos que não se enquadram no frame. Em seguida, é feita uma comparação desses achados com o que dicionários apresentam. Ambas as tarefas contribuem para clarificar as delimitações do próprio frame, de modo que o anotador pode caracterizá-lo mais precisamente.

A definição do frame pode ser considerada uma “apresentação esquemática de um tipo de situação que subjaz ao significado de uma palavra (ou dos membros de grupos de palavras) juntamente com papéis de participantes nomeados ou aspectos da situação” (FILLMORE ET AL., 2003, p. 305).¹⁹ A respeito dos últimos, aqui abordados como EFs, embora possam ser considerados instâncias de papéis semânticos mais amplos – tais como Agente e Paciente – Fillmore et al. alegam que, ao usar papéis micro temáticos é possível oferecer definições precisas aos EFs sem que seja necessário decidir como encaixá-los em um inventário pré-determinado e reduzido de papéis semânticos (2003, p. 305).

Os EFs podem ser divididos em nucleares, periféricos e extra-temáticos. Como mencionado, enquanto os nucleares retratam aqueles elementos indispensáveis à composição da cena, os periféricos trazem informação adicional, de modo que tais elementos não são responsáveis por caracterizar um dado frame. Por fim, os extra-temáticos trazem à tona outro estado de coisas, de modo que os termos etiquetados como EFs extra-temáticos são comumente ULs pertencentes a outro frame. Além disso, como lembram Fillmore et al. (2003, p. 319), EFs não nucleares não podem ser sujeito ou objeto de um verbo gramaticalmente e amiúde são sintagmas preposicionados ou advérbios.

Quando da anotação de sentenças, se um EF nuclear não consta, isso é registrado – a não ser que haja alguma informação explícita no frame que

¹⁹ ... a schematic presentation of a situation type that underlies the meaning of a word (or of the members of sets of words) along with named participant roles or aspects of the situation...

permita seu apagamento, condicionado à presença de outro EF nuclear²⁰ –, ao passo que não há nenhuma sinalização que indique EFs não nucleares ausentes. Assim, os EFs nucleares que não se manifestam como material lexical ou sintagmático em uma dada sentença são chamados de instanciações nulas, as quais se dividem em três tipos: construcional, definida e indefinida. É válido destacar que o fato de um EF nuclear não se manifestar linguisticamente não é sinônimo de que ele não tenha sido utilizado pelo falante na compreensão, como veremos a seguir.

Fillmore et. al. (2003, p. 320-321) esclarecem que a Instanciação Nula Construcional (CNI) é licenciada por uma construção gramatical – como a omissão do agente em sentenças passivas. A CNI existe como um atalho para sinalizar situações desse tipo enquanto ainda não foi criada uma construção que licencie tal realização, de modo que, num cenário em que todas as construções relevantes de uma língua sejam registradas em um repertório e suas instanciações anotadas nas sentenças, as CNIs perdem sua utilidade. Por sua vez, na Instanciação Nula Definida (DNI), é preciso que o referente do elemento ausente seja passível de ser compreendido através do contexto linguístico ou discursivo. Por fim, a Instanciação Nula Indefinida (INI) diz respeito à elipse não de referentes específicos, mas de categorias inteiras de referências, tais como os objetos faltantes de determinados verbos transitivos – como “comer”, “assar” e “beber” – usados de modo intransitivo.

Todavia, em função de disparidades entre a língua portuguesa e o inglês norte-americano, algumas alterações foram realizadas quanto à classificação das instanciações nulas. Torrent et al. (no prelo) alegam que a principal diferença entre a DNI e a INI é, mesmo, o fato de que a primeira apresenta uma anáfora zero em que se pode identificar um referente, ao passo que a segunda é uma omissão existencial que não permite a identificação de um referente no contexto. Porém, a versão indefinida, para o inglês, é tratada de modo especial como uma propriedade da UL. O exemplo (15) demonstra a questão:

- (15) *Would you like to try some of this delicious cake?*
 a. *No, thanks, I already ate.*
 b. *No, thanks, I already ate it.*

²⁰A relação de *Excludes*, modelada entre EFs de um mesmo frame, define que, na presença de um dado EF, outro EF não possa ocorrer. Por outro lado, os Core Sets são conjuntos de EFs que podem estar ausentes numa sentença, desde que ao menos um deles se manifeste.

c. Não, obrigado, eu já comi.
No, thanks, I already ate. / No, thanks, I already ate it.

A pergunta em (15) traz um exemplo de INI com o verbo “eat” (comer) na resposta (a), a qual denota que o falante já comeu outro alimento antes. Por outro lado, se o falante tivesse comido do bolo previamente, a resposta seria a que aparece em (b). Contudo, em português brasileiro o mesmo não ocorre, visto que a resposta (c) seria adequada para representar ambos os cenários anteriores, isto é, aquele em que o falante já comeu outro alimento bem como aquele em que o falante já comeu do bolo.

Torrent et al. (no prelo) explicam que uma vez que o português brasileiro permite, de maneira geral, omissões de objeto direto para ambas as situações – referência anafórica e existencial –, considera-se somente o *status* informacional do EF omitido. Assim, na FrameNet Brasil, ocorrerá um caso de DNI quando for possível identificar um referente específico; quando não o for, um rótulo de INI deverá ser usado.

Ressalta-se que na FrameNet anotam-se apenas as partes da sentença que são conectadas sintaticamente à UL alvo, uma vez que o objetivo do trabalho é realizar um mapeamento entre sintaxe e semântica (TORRENT & ELLSWORTH, 2013, p.53).

Existe, ainda, um Editor de UL, por meio do qual é possível detalhar, a respeito da UL em questão, sua forma de dicionário, sua classe de palavra, seu significado e sua composição formal (FILLMORE ET AL., 2003, p.313). Conforme Torrent & Ellsworth (2013, p.45) explicam, as ULs recebem uma definição que pode ser escrita pelo anotador ou obtida a partir de um dicionário. As ULs alvo da FrameNet são, majoritariamente, verbos, substantivos, adjetivos, advérbios e preposições, além de determinadas expressões polilexêmicas, algumas conjunções, numerais e pronomes indefinidos.

As sentenças em que as ULs se manifestam são analisadas localizando-se os EFs nelas presentes e qualificando-os quanto à Função Gramatical (FG) e ao Tipo Sintagmático (TS) que apresentam, conforme observa-se na Figura 12. Os EFs são anotados considerando-se constituintes inteiros, e não apenas o núcleo do sintagma. Fillmore et. al. (2003, p.319) apontam que, na Berkeley

extraídas para uma dada UL. Logo, não há embasamento efetivo para se afirmar quais são as possibilidades combinatórias mais frequentes.

Há, ainda, um modo de sinalizar generalizações semânticas de maneira mais econômica. Conforme Fillmore et al. (2003, p.310) explicam, após realizar a definição do frame e dos EFs, é possível apontar relações entre o frame atual e outros frames por meio do Editor de Relação. Atualmente, a FrameNet trabalha com nove relações, quais sejam, Herança, Uso, Perspectiva, Subframe, Precedência, Incoativo de, Causativo de, Metáfora e Veja também.

Um dos benefícios para a criação de relações entre frames, segundo alegam Ruppenhofer et al. (2016, p.79), é o fato de que o significado de um frame abstrato pode ser mais facilmente compreendido por meio de sua relação com um frame mais concreto. Além disso, com a rede de frames, contribui-se, ainda, para a robustez da análise, visto que frames que estão separados podem ser relacionados àqueles que são similares semanticamente.

As relações entre frames são assimétricas, de modo que um frame, o mais abstrato – ou menos dependente – é chamado, genericamente, de Super frame, ao passo que o outro, menos abstrato – ou mais dependente – é chamado de Sub frame. Em cada relação, contudo, Super e Sub frames recebem nomes específicos (RUPPENHOFER ET AL., 2016, p.79), como se observa no Quadro 1.

Ruppenhofer et al. (2016, p.80) assinalam que a relação de Herança é considerada a mais forte e destina-se a modelar fatos de aceção intuitiva. Nessa relação, o que for verdadeiro para o frame pai deve corresponder a algo mais específico no frame filho, como se observa no exemplo a seguir: o EF AGENTE no frame de *Causar_mover* é aquele cuja ação causa o movimento de um TEMA. No frame de *Passar*, 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. Assim, o movimento realizado pelo EF AGENTE é especificado em relação ao mesmo EF em *Causar_Mover*. Logo, o frame de *Passar* herda de *Causar_Mover*.

Nos casos em que um frame filho pressupõe o frame pai como pano de fundo (“background”), se estabelece uma relação de Uso entre eles, a qual é empregada, segundo Ruppenhofer et al. (2016, p.83), quase exclusivamente para casos em que uma parte da cena evocada pelo frame filho refere-se ao

frame pai. 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.

Relação	Super frame	Sub frame
Herança	Pai	Filho
Uso	Pai	Filho
Perspectiva	Neutro	Perspectivado
Subframe	Complexo	Componente
Precedência	Anterior	Posterior
Incoativo de	Incoativo	Estativo
Causativo de	Causativo	Incoativo/ Estativo
Metáfora	Fonte	Alvo
Veja Também	Entrada Principal	Entrada Referente

Quadro 1: Tipos de relação entre frames (RUPPENHOFER ET AL., 2016, p.80)

A relação de *Perspectiva*, por sua vez, indica a presença de, pelo menos, dois pontos de vista que podem ser assumidos a partir de um frame neutro (RUPPENHOFER ET AL., 2016, p.82). Geralmente, o frame neutro é não-lexical e não-perspectivado. Os frames de *Comércio_comprar* e *Comércio_vender* se relacionam com o frame de *Comércio_transferência_de_mercadoria* – frame neutro, não-lexical – pela relação de *Perspectiva*, a saber, a do *COMPRADOR* e a do *VENDEDOR*, respectivamente.

Conforme Ruppenhofer et al. (2016, p.83), alguns frames são complexos na medida em que se referem a sequências de estados e transições, cada uma das quais podendo ser separadamente descrita 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`.

Já a relação de Precedência ocorre somente entre dois frames componentes de um único frame complexo. Essa relação captura a ordenação temporal de subeventos dentro de um evento complexo (PETRUCK & DE 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 – são subframes do frame de `Ciclo_de_dormir_e_acordar` – evento complexo.

Os frames estativos, incoativos e causativos se relacionarão através das relações de Incoativo de e Causativo de, como em “Ana e Paula secaram o pano”, do frame `Causar_estar_seco`, o qual é Causativo do frame de `Estar_seco`, como em “O pano estava seco”.

A relação de Metáfora, a mais nova implementada, estabelece conexão entre um frame Fonte e um frame Alvo em que “muitas ou todas as ULs no frame Alvo são entendidas ao menos parcialmente em termos do frame Fonte” (RUPPENHOFER ET AL., 2016, p.85).

Finalmente, a relação de Veja também é empregada nos casos em que grupos de frames são similares e deveriam ser comparados e contrastados (RUPPENHOFER ET AL., 2016, p.85). Trata-se de uma metarrelação, visto que ela é empregada para solucionar possíveis dúvidas do consulente.

A FrameNet apresenta, ainda, outro modo de anotação além do lexicográfico, qual seja, a anotação de texto corrido. Nessa metodologia, ao se adotar um determinado texto, realiza-se a análise de todas as ULs nele presentes, de maneira que diferentes frames serão evocados e anotados.

Conforme sintetiza Salomão, os objetivos principais da Berkeley FrameNet são:

- (i) descrever Unidades Lexicais (ULs) a partir dos frames que evocam, frames que são identificados em termos dos Elementos de Frame (EFs) que crucialmente os constituem;
- (ii) validar tais descrições através de pesquisa em *corpora*;
- (iii) determinar as possibilidades combinatórias das ULs pesquisadas, anotando, nas sentenças extraídas dos *corpora*, em camadas, quais são os

EFs expressos por cada sintagma, bem como seu Tipo Sintagmático (TP) e Função Gramatical (FG);
 (iv) disponibilizar os resultados da anotação na forma de entradas lexicais, as quais sumarizam os possíveis padrões de valência de cada UL;
 (v) e definir as relações entre frames, mostrando, por exemplo, em que medida um frame é uma elaboração de outro ou se depende de outro, relações representadas via ferramenta FrameGrapher, como se pode verificar no site do projeto FrameNet (SALOMÃO, 2013, p.12).

A partir de tais princípios, a Berkeley FrameNet cresceu e projetos irmãos foram desenvolvidos para outros idiomas. Atualmente, há ramos consolidados para o alemão (BOAS, 2002), o chinês (YOU & LIU, 2005), o espanhol (SUBIRATS & PETRUCK, 2003), o japonês (OHARA ET AL., 2004), o sueco (BORIN ET AL., 2010) e o português (SALOMÃO, 2009b). Torrent e Ellsworth (2013, p.44) apontam que as diversas framenets costumam utilizar os mesmos frames definidos inicialmente em Berkeley, o que não ocorre com as ULs, uma vez que cada língua apresenta suas próprias unidades. As valências sintáticas, compostas de Tipos Sintagmáticos e Funções Gramaticais, também podem variar entre os diversos idiomas, ao passo que as valências semânticas são definidas em função dos EFs, os quais dependem do frame e, portanto, tendem a encontrar mais correspondências entre as línguas, apesar de isso não ser necessariamente verdade em todos os casos.

Embora o planejamento inicial da FrameNet tivesse por objetivo, como afirmado anteriormente, caracterizar as principais propriedades distribucionais de palavras do inglês através da apresentação de seus padrões de valência, com o tempo, tornou-se evidente que determinados aspectos não eram contemplados pela anotação lexicográfica. Nesse contexto, nasce um novo projeto para lidar com tais casos: o Constructicon, o qual será apresentado na próxima seção.

2.3.2 *Constructica*

2.3.2.1 Fundamentos

Com o intuito de reconhecer e catalogar as construções gramaticais do inglês, surge, em 2008, como projeto apêndice à FrameNet, o Constructicon. Como explicam Fillmore, Lee-Goldman e Rhomieux (2012), o novo empreendimento receberia a função de descrever características gramaticais e

informação semântica contidas nas construções sob análise, disponibilizando sentenças anotadas para cada construção com fins de exemplificação. Esse trabalho seria implementado utilizando-se as ferramentas já desenvolvidas para o recurso lexical da FrameNet.

Embora uma construção possa ser descrita na forma de uma matriz de valores de atributos ou informalmente em forma de prosa, a anotação repousa sobre o construto, isto é, uma porção da língua licenciada por uma construção (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012). Assim, “cada anotação captura as propriedades de um construto particular com relação a uma construção particular que o licencia” (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012, p. 321).²³ Os autores ilustram o funcionamento da anotação por meio da construção *Month_plus_date* (*Mês_mais_data*), a qual licencia expressões como *June 14th* (14 de junho) e *May 2nd* (2 de maio). Para anotar tal construção no construto *June 14th*, é preciso identificar as duas partes necessárias para esse sintagma: *June* e *14th*, os quais correspondem às categorias *MONTH-NAME* (*NOME_DE_MÊS*) e *ORDINAL-NUMBER* (*NÚMERO_ORDINAL*), respectivamente. Ao mesmo tempo, outras construções que podem estar envolvidas na constituição do referido sintagma, como a construção que habilita números ordinais, por exemplo, são ignoradas uma vez que cada anotação recruta apenas uma construção por vez.

Fillmore, Lee-Goldman e Rhomieux (2012, p. 310) apontam que a anotação construcional, adaptada da lexicográfica, precisa indicar três informações importantes a respeito das construções: (i) os trechos da língua que se configuram como instâncias de determinadas construções – na sentença a seguir, o fragmento em negrito indica o que instancia a construção de *Dativo_com_Infinitivo_Suficiência*: “Ela deu **dinheiro pra mim viajar**” –; (ii) os segmentos desses trechos que são ditados pelos requisitos da construção – trechos separados por colchetes: “Ela deu [**dinheiro**][**pra mim viajar**]” –; e (iii) elementos do contexto cujas propriedades são selecionadas pela instância da construção – segmento entre chaves: “{Ela deu}[**dinheiro**][**pra mim viajar**]”.

²³ ... each annotation captures the properties of a particular construct with respect to a particular construction that licenses it.

A primeira informação a ser elencada por meio da anotação indica o signo mãe do construto, ao passo que seus componentes, identificados na segunda etapa de anotação, são os signos filhos, denominados Elementos da Construção (ECs) na presente metodologia. Desse modo, no exemplo acima, o signo Mãe corresponde a “dinheiro pra mim viajar”, o qual apresenta dois ECs, *DINHEIRO* e *PRA_MIM_VIAJAR*.

No Constructicon, a construção, tal como um frame, recebe uma descrição em prosa acerca de suas propriedades, a qual pode apontar frames evocados. Retomando a construção *Month_Plus_Date*, anteriormente aludida, em sua descrição haverá menção aos frames relevantes relativos a datas e meses, por exemplo, e será atribuído ao construto a informação de que ele designa uma unidade de "dia" (e não de semana, mês, ano etc. (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012, p. 322).

No caso de uma construção evocar um frame, cada EC é definido em prosa e, às vezes, através de seu próprio nome, segundo o EF do frame evocado. Logo, como os autores exemplificam (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012, p. 322), a construção *Verb-way*, a qual combina uma classe de verbos – descrita por Goldberg (1995) – a um sintagma nominal possessivo nucleado pelo termo “way” para criar um verbo polilexêmico que evoca o frame de *Motion* (Movimento), é representada conforme (16).

(16) *Hopefully* [*Theme* *he*]’ll *make his way* [*Goal* *to our location*].

Com esperança, ele conseguirá chegar à nossa localização.

Observam-se, na sentença (16), os ECs *TEMA* e *ALVO* conectados ao material linguístico que os instancia: *he* e *to our location*, respectivamente. Fillmore, Lee-Goldman e Rhomieux (2012, p.322) afirmam que essa forma de anotação construcional permite observar os EFs de *Motion* presentes em qualquer sentença exemplar da construção *Verb-way*. Entretanto, como se mostrará nesta tese, não há uma relação modelada na base da Berkeley FrameNet que conecte construções a frames. Ademais, alegam que a inserção de tais informações condiz com a anotação lexical da FrameNet, visto que há camadas para assinalar inclusive Função Gramatical e Tipo Sintagmático dos

elementos – para além de uma harmonia teórica, havia necessidade de adaptação do software utilizado na empreitada lexical à construcional.

Em analogia com a anotação lexicográfica, os pesquisadores (2012, p.323) explicam também que, no lugar da UL alvo, há um construto potencialmente polilexêmico. A construção é delimitada pelo rótulo TSConstruto (CstrPT), o qual indica o Tipo Sintagmático (TS) do signo mãe. Em seguida, os constituintes internos do construto são identificados na camada dos ECs. Na construção *Month_Plus_Date*, por exemplo, o TSConstruto será assinalado como um sintagma nominal (“*June 14th*”) e a cada um dos ECs, *MONTH_NAME* (“*June*”) e *ORDINAL_NUMBER* (“*14th*”), será atribuído o TS sintagma nominal.

Na Figura 13, observa-se um exemplo de anotação construcional proporcionado por Fillmore, Lee-Goldman e Rhomieux (2012, p.324), com o construto “*Constance squeezed her way down the platform*”. Como os autores salientam, o TSConstruto (CstrPT) – “*squeezed her way*” – é um sintagma verbal finito; e o sintagma “*her way*” é identificado como o Elemento Evocador da Construção (EEC), isto é, o elemento responsável por sinalizar material lexicalmente especificado que serve de gatilho à evocação da construção no processo de interpretação, se houver.

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																			C	E	E															
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 13: Anotação de uma instância da construção *Verb-way* (FILLMORE, LEE-GOLDMAN & RHOMIEUX, 2012, p. 324)

Na primeira camada, CE (*Construction Element*), os ECs são identificados, como *TEMA* (*THEME*), *VERBO_TRANSITIVO_MEIO* (*TRANSITIVE_MEANS_VERB*), *EXPRESSÃO_WAY* (*WAY_EXPRESSION*) e *CAMINHO* (*PATH*). Em seguida, a camada GF (*Grammatical Function*) aponta a FG dos ECs localizados, como Externo (Ext), para o *TEMA*, e Dependente (Dep), para o *CAMINHO*. A camada PT (*Phrase Type*) indica os TS dos ECs, SN (NP) para o *TEMA*, SVfin (VPfin) para o *VERBO_TRANSITIVO_MEIO*, SN para a *EXPRESSÃO_WAY* e SP (PP) para o *CAMINHO*. Em CEE (*Construction Evoking Element*), o EEC é registrado: a *EXPRESSÃO_WAY*. Por fim, o CstrPT é sinalizado: SVfin (VPfin), um sintagma verbal finito.

Em suma, no Constructicon de Berkeley, cada construção apresenta as seguintes informações:

1. Uma fórmula entre parênteses que oferece nomes mnemônicos para os constituintes mãe e filho;
2. Um nome mnemônico para a construção;
3. Uma descrição informal das propriedades do constituinte mãe;
4. Descrições informais de cada um dos constituintes filhos;
5. Uma interpretação da maneira pela qual as propriedades dos filhos participam nas propriedades detalhadas do signo resultante, em termos de sintaxe, semântica, pragmática e contexto. (FILLMORE, LEE-GOLDMAN E RHOMIEUX, 2012, p.331)²⁴

Torrent et al. (no prelo) argumentam que desenvolver um Constructicon é desafiador na medida que é preciso representar computacionalmente restrições e possibilidades das construções, isto é, os modelos computacionais desse recurso devem contemplar a constituência e a unificação em construções. No Constructicon de Berkeley, a primeira propriedade é abordada por meio da identificação dos ECs. A segunda propriedade, por sua vez, é tratada através da anotação de instâncias da construção em camadas, ao alinhá-las – contemplando-se FG e TS descritos – aos ECs (FILLMORE ET AL., 2012, p. 321-324).

A partir do Constructicon de Berkeley, outras iniciativas semelhantes estão em implementação avançada para o Sueco (LYNGFELT ET AL., 2012) e para o Português Brasileiro (TORRENT ET AL., 2014), enquanto aquelas para o

²⁴ 1. a bracketing formula offering mnemonic names for the mother and daughter constituents; 2. a mnemonic name for the construction; 3. an informal description of the properties of the mother constituent; 4. informal descriptions of each of the daughter constituents; 5. an interpretation of the manner in which properties of the daughters participate in the detailed properties of the resulting sign, in terms of syntax, semantics, pragmatics, and context..

Japonês (OHARA ET AL., 2013) e o Alemão (ZIEM & BOAS, 2017) encontram-se em fase mais inicial.

Na seção seguinte, apresentamos as características próprias do Constructicon da FrameNet Brasil, no qual foram implementadas as relações modeladas por esta tese.

2.3.2.2 Constructicon da FrameNet Brasil

Em 2010, o Constructicon do português brasileiro foi iniciado a partir de análises das construções da Família Para Infinitivo²⁵ (TORRENT, 2015), a qual será apresentada posteriormente. É importante ressaltar que o Constructicon da FN-Br modela as construções de maneira diferente daquela empregada na versão original, como veremos a seguir.

Conforme se observa em Lage (2013), foram estabelecidas políticas de anotação que oferecem critérios objetivos para a criação de construções. Para que existisse um princípio claro que definisse quais construções deveriam ser anotadas no Constructicon e quais poderiam ser tratadas via padrão de valência no Lexicon, três critérios foram instituídos. O primeiro critério concerne à existência de algum material linguístico lexicalmente especificado na potencial construção. Como a anotação lexicográfica na FrameNet sempre parte de uma UL, é uma condição *sine qua non* que haja algum item lexical específico. Eis o primeiro critério:

Critério 1: Onde X é um material lexicalmente especificado, há X na potencial construção?

Se a resposta a tal pergunta é negativa, sabe-se que a construção alvo não poderá ser anotada no Lexicon e, portanto, será direcionada ao Constructicon. Ao contrário, se a resposta for positiva, deve-se prosseguir para o segundo critério, o qual examina se a potencial construção evoca ou não um frame. As unidades básicas de análise na anotação lexicográfica, conforme visto,

²⁵ Adotando as mesmas diretrizes que o Projeto Beyond the Core, iniciado pela FrameNet de Berkeley, o Constructicon da FrameNet Brasil começou suas representações computacionais por construções que já haviam sido estudadas por gramáticos. As primeiras a serem analisadas foram aquelas pesquisadas por linguistas da Universidade Federal de Juiz de Fora.

são as ULs, pareamentos de um lema a um frame. Logo, é condição *sine qua non* que o elemento sob análise evoque um frame.

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

Se a potencial construção não evoca frame, não poderá ser abordada como uma UL; deverá, portanto, receber tratamento construcional. Contudo, se ela for capaz de evocar frame, parte-se para o terceiro e último critério. Nele, verifica-se a possibilidade de a potencial construção evocar o mesmo frame em diferentes padrões de valência. Se isso for possível, significa que o frame foi evocado pela unidade alvo, o que conduz a um tratamento lexicográfico, uma vez que a anotação da UL é suficiente. Todavia, a resposta negativa direciona a anotação ao Constructicon.

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

Além dos três critérios mencionados, Torrent et al. (no prelo) assinalam que outra importante diferença pode ser notada na metodologia do Constructicon do português brasileiro, o qual cria ECs baseados em aspectos formais, em vez de considerar traços funcionais/semânticos. Para contemplar essas propriedades da construção, uma relação de evocação entre construções e frames foi desenvolvida no escopo desta tese, conforme a seção 5 mostrará. O Constructicon brasileiro difere-se, ainda, daquele do sueco na medida que, no último, as construções têm papéis semânticos genéricos, tais como, Agente e Paciente (TORRENT ET AL., no prelo).

Há que se ressaltar, ainda, que uma das principais finalidades do Constructicon é servir de base de dados para alimentar um analisador sintático-semântico, o CARMA²⁶ (MATOS ET AL., 2017). Logo, além de prover uma interface legível por humanos para o recurso, é necessário também que as características da construção modelada sejam legíveis por máquina.

Nesse contexto, o processo de criação de uma construção no Constructicon da FrameNet Brasil se inicia por sua definição e a dos ECs. Para

²⁶ O CARMA – *Constructional Analyzer using Relations among Multiple Attribute-Value Matrices* – é o analisador construcional que vem sendo desenvolvido pela FN-Br.

ilustrar tal tarefa, será utilizada a construção *Transitiva Direta Ativa*, que começa a ser inserida a partir da indicação de três informações, conforme mostra a Figura 14.

Figura 14: Edição de entrada

Em “*entry*”, é colocado o nome interlingual da construção – *cxn_act_trans_dir* –, o qual permite a sua identificação. O termo “*cxn*” é uma abreviação para “construção”, ao passo que o restante do nome remete à especificação da construção (*Transitiva Direta Ativa*). Em seguida, deve-se apontar se a construção é abstrata – ou seja, se licencia construtos na língua [- abstrata], ou se apenas estrutura a rede hierárquica de construções [+ abstrata] –, e se está ativa na base de dados que alimenta o CARMA, que serão marcados para esta construção em negativo e em positivo, respectivamente, indicando que essa construção de fato licencia construtos em PB e que foi revisada e liberada para uso no parser. Após salvar essas informações, uma nova tela se abrirá, representada na Figura 15.

Lang	Name	Description
en	Active_Transitive	[NP[V NP]] construction type. This construction ex...
es	cxn_act_trans_dir	cxn_act_trans_dir...
fr	cxn_act_trans_dir	cxn_act_trans_dir...
pt	Transitiva Direta Ativa	Tipo de construção SN_V_SN. Essa construção exhibe ...

Figura 15: Edição de entrada por idioma

Nessa nova tela, é possível criar entradas legíveis por usuários humanos em diversos idiomas. Clicando em “pt”, o anotador é direcionado a uma nova tela – Figura 16.

The screenshot shows a dialog box titled "cxn_act_trans_dir [pt]". At the top left is a "Save" button. Below it are three input fields: "Name" containing "Transitiva Direta Ativa", "Description" containing "Tipo de construção SN_V_SN. Essa construção exibe um argumento externo #Sujeito e um #Predicado. O argumento interno é um objeto direto.", and "Nick" containing "cxn_act_trans_dir".

Figura 16: Definição da construção

Na Figura 16, insere-se o nome da construção e sua definição, que deve ser pautada por aspectos sintáticos. Abaixo, em “nick”, permanece sinalizado o ID da construção. Feito isso, a construção está criada, sendo o próximo passo a criação e a definição dos ECs.

The screenshot shows a dialog box titled "CxnElement: ce_predicate_act_trans_dir []". At the top left are "Close" and "Save" buttons. Below are several fields and controls: "Entry:" with "ce_predicate_act_trans_dir", "Color" with a dropdown menu showing "white_blue", and three toggle buttons: "Optional" (set to No), "Head" (set to Yes), and "Multiple" (set to No).

Figura 17: Edição de Elemento da Construção Predicado

A Figura 17 apresenta a tela de edição do EC, onde primeiro se coloca o nome interlingual e, em seguida, a cor que será usada para identificar tal EC. A primeira parte da entrada, “ce_predicate”, refere-se ao fato de que se trata de um EC cujo nome é “predicado”, enquanto o restante faz menção à construção

à qual pertence tal EC – “act_trans_dir”. A seguir, o anotador deve selecionar se tal EC é opcional, se é nuclear e se é múltiplo. Neste caso, o EC precisa constar na sentença, por isso ele não é opcional; é nuclear, uma vez que os traços do verbo que compõe o predicado serão assumidos pela construção como um todo; mas não é múltiplo, uma vez que não há a possibilidade de vários *PREDICADOS* recursivamente concatenados nessa construção, sem a presença de uma construção de *Coordenação*.

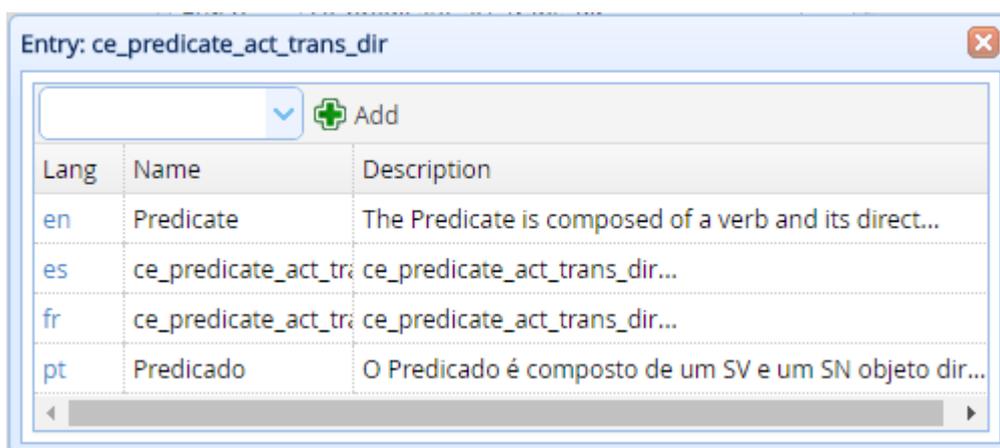


Figura 18: Edição de Elemento da Construção por idioma

Na Figura 18, observa-se a possibilidade de inserir o EC em um dos quatro idiomas disponíveis. Ao clicar em “pt”, uma nova tela se abre, conforme a Figura 19 exhibe.

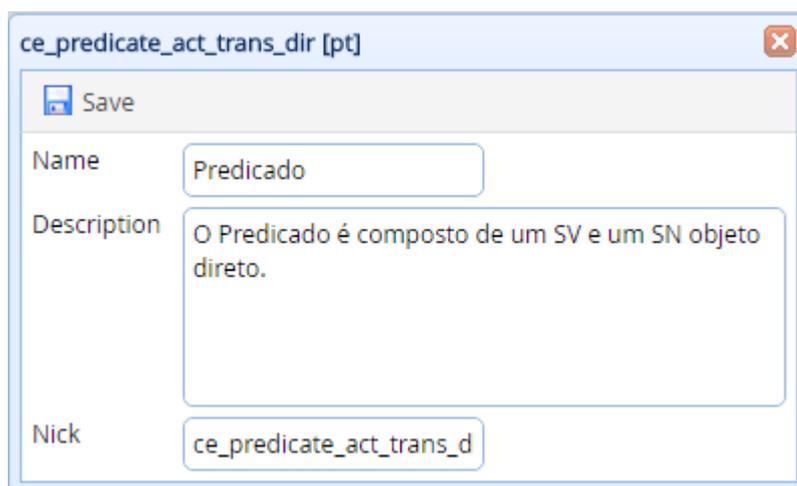


Figura 19: Definição do Elemento da Construção Predicado

O anotador pode, então, salvar o nome do EC, *PREDICADO*, bem como defini-lo. O mesmo procedimento é realizado para se registrar o outro EC desta construção, o *SUJEITO*. Assim, está criada uma construção composta de seus ECs, o que produzirá a tela da Figura 20 para o usuário do Constructicon.

Transitiva Direta Ativa [cxn_act_trans_dir]

Definição	
Tipo de construção SN_V_SN. Essa construção exibe um argumento externo Sujeito e um Predicado . O argumento interno é um objeto direto.	
Exemplo(s)	
Elementos da Construção	
Predicado [ce_predicate_act_trans_dir]	O Predicado é composto de um SV e um SN objeto direto.
Sujeito [ce_subject_act_trans_dir]	O sujeito é um SN.
Relações	
Evoca	Ação_transitiva
Herdado por	Objeto_interdito
Herda de	Sujeito_predicado

Figura 20: Construção *Transitiva_Direta_Ativa*

Nota-se que a tela é intitulada com o nome da construção, seguida de seu nome interlingual. A seguir, apresenta-se sua definição, os exemplos de construtos licenciados pela *Transitiva_Direta_Ativa* e seus ECs. Por fim, há uma lista que aponta as relações em que se insere esta construção.

Através dos passos mostrados para o estabelecimento de uma construção no Constructicon, nota-se que as informações disponibilizadas são legíveis para usuários humanos. Entretanto, lembrando o direcionamento desse recurso a tarefas em NLU – como, por exemplo, o já citado CARMA –, é imprescindível que tais informações sejam formalizadas de modo que seja possível a leitura por máquinas.

Nesse sentido, as restrições são empregadas como instrumento capaz de sinalizar uma série de informações a serem processadas automaticamente. Há, portanto, três restrições disponibilizadas no Editor de Restrições: a *CE_Construction*, a *CE_before* e a *CE_meets*.

A primeira restrição, *CE_Construction*, é empregada para determinar se o signo filho é licenciado por outra construção do Constructicon. Em outras palavras, se um EC for composto por uma construção já definida no

Constructicon, essa informação poderá ser assinalada. Isso possibilita ao sistema economia, na medida em que todas as restrições da construção mãe são automaticamente incorporadas ao EC. No caso da construção *Transitiva_Direta_Ativa*, o EC *PREDICADO* e o EC *SUJEITO* têm essa restrição, já que aquele é licenciado pela construção do *Sintagma_Verbal_com_Complemento*, e este pelo *Sintagma_Nominal_Determinado*.

A segunda restrição é a *CE_before*, a qual determina se um dado EC deve vir antes de outro na sentença. Essa restrição é assinalada ao EC *SUJEITO*, que deve preceder o *PREDICADO*.

A terceira e última restrição já implementada é a *CE_meets*, que indica que um EC deve vir antes de outro e que ambos devem ser adjacentes, ou seja, não pode haver nenhum material interveniente entre eles. No caso da construção *Transitiva_Direta_Ativa* essa restrição não se aplica.

Aplicadas as restrições, estas são registradas na interface, de modo a permitir que o analista veja as propriedades das construções constantes do Constructicon, conforme mostra a Figura 21.

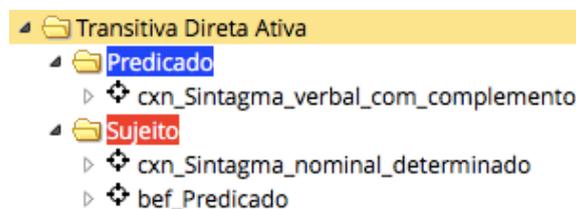


Figura 21: Restrições aplicadas à construção *Transitiva_Direta_Ativa*

Através desse aparato, as construções são definidas no Constructicon do PB. Entretanto, ao lidar com algumas construções, foi observado que uma série de propriedades ainda não podia ser capturada pelas restrições já definidas no recurso construcional, conduzindo a pesquisa por um levantamento acerca do que seria necessário implementar a fim de garantir uma descrição completa e formalizada. Esse é o assunto discutido nesta tese.

Antes de passarmos a apresentar as propostas implementadas, no âmbito desta tese, para o Constructicon do PB, aprofundaremos, no capítulo 3, a

discussão sobre as relações entre construções e entre estas e os frames, conforme definidas pelos modelos teóricos construcionistas.

3 RELAÇÕES ENTRE CONSTRUÇÕES E ENTRE FRAMES E CONSTRUÇÕES

Este capítulo discute as versões da Gramática das Construções, analisadas no primeiro capítulo, pelo viés da herança construcional e das implementações que lidam com frames e construções.

3.1 Redes Construcionais: Herança Completa versus Herança Normal

Conforme apontado previamente, Fillmore assinala que uma gramática é um repertório de construções. Contudo, “em vez de ser apenas uma coleção não estruturada, as construções gramaticais em uma língua formam uma rede conectada por links de herança” (FILLMORE, 1999, p.115)²⁷. Esses links são o tópico desta seção.

Goldberg (1995, p.67) concorda com Fillmore ao apontar que, ao contrário de um conjunto desestruturado, encontram-se generalizações sistemáticas perpassando as construções. Para este linguista, as generalizações gramaticais são reconhecidas na medida em que se pode observar de que forma certas construções contêm ou são elaborações de outras construções; ele explica: “se a construção C herda da construção D, então C compartilha todas as condições de D enquanto adiciona algumas próprias” (FILLMORE, 1999, p.115).²⁸

Foi nas Ciências da Computação que a herança passou a ser usada como modo de capturar generalizações. As informações são armazenadas de forma eficiente e tornadas facilmente modificáveis, ao se postular hierarquias nas quais os níveis mais baixos herdam informação dos níveis mais altos, como afirma Goldberg (1995, p.72).

Nesse contexto, uma descrição completa, por exemplo, da construção do inglês *SAI* (*Subject-Auxiliary_Inversion* – *Inversão_de_Sujeito_e_Auxiliar*) deve conter todas as propriedades de cada construção que a antecede em uma hierarquia, argumenta Fillmore (1999, p.121). Como a *SAI* herda das construções *VHP* (*Verb-Headed_Phrase* – *Sintagma Nucleado por Verbo*), *LHP* (*LexicallyHeadedPhrase* –

²⁷(...) instead of being just an unstructured collection, the grammatical constructions in a language form a network connected by links of inheritance.

²⁸ If construction C inherits construction D, then C shares all of the conditions of D while adding some of its own.

Sintagma Lexicalmente Nucleado), *VR* (*Valence Raising - Alçamento de Valência*), *LTP* (*Lexically Tagged Phrase - Sintagma Marcado Lexicalmente*) e *HFC* (*Head Feature Construction - Construção de Traços de Núcleo*), sua plena descrição precisa considerar as características dessas cinco construções de que ela herda. De *VHP*, ela herda o fato de o núcleo do signo filha que requer um argumento sujeito ser da categoria Verbo, enquanto, a herança de *LHP* traz a informação mais genérica de que, em inglês, um dos subtipos possíveis de um núcleo lexical é um Verbo. Já as construções *LTP* e *HFC*, dão conta da tipagem da *SAI* como uma construção verbal e da projeção dos traços do núcleo verbal para a sintaxe da construção como um todo, ou seja: toda a sentença que traz a inversão de sujeito-auxiliar se comporta segundo propriedades do verbo que lhe serve de núcleo. Por fim, *VR* faz com que a valência especificada para o verbo que nucleia a construção seja adotada pela construção como um todo.

Note-se que estas propriedades não são exclusivas da *SAI*, mas de qualquer construção que tenha um núcleo lexical. É nesse sentido que Fried e Ostman (2004, p.72) explicam que as relações de herança fornecem um mecanismo para representar a gramática como uma rede de padrões gramaticais relacionados e muitas vezes sobrepostos. Esclarecem, também, que construções que compartilham um traço particular compõem *clusters* de generalizações mutuamente relacionadas, indo da mais abstrata e sem restrições à que possui mais restrições (FRIED & OSTMAN, 2004, p.72).

Há dois possíveis caminhos de herança: a única e a múltipla. Enquanto na herança única uma dada construção em uma hierarquia pode herdar apenas de uma construção dominante, na múltipla é possível herdar de mais de uma construção dominante. Em outros termos, na primeira, cada construção filha pode ter apenas uma construção mãe, ao passo que na segunda, é possível uma construção filha ter mais de uma construção mãe. No Constructicon, a herança múltipla é adotada, assim como o fazem os demais modelos de Gramática das Construções. A construção de *Dativo_com_Infinitivo*, por exemplo, herda ao mesmo tempo da *Oração_relativa*, por ser um tipo de relativização, e da *Para_infinitivo*, devido às suas propriedades formais. Do mesmo modo, exemplifica a herança múltipla a supracitada construção *SAI*, a qual apresenta cinco diferentes construções mãe.

A herança se divide, também, entre o modo normal e o completo. Segundo Goldberg (1995, p.73), no primeiro – empregado na Gramática Cognitivista das Construções –, o qual permite subregularidades e exceções, a informação é herdada dos nós dominantes de forma transitiva, desde que essa informação não entre em conflito com a informação especificada por nós mais baixos na hierarquia de herança. Trata-se de constatar generalizações parciais, dado que a construção filha possui determinadas propriedades da construção mãe, mas pode violar algumas delas. Por sua vez, o modo completo é projetado para capturar relações e restrições taxonômicas. Neste modo – implementado pela BCG –, toda informação específica para cada nó que domina, direta ou indiretamente, um determinado nó é herdada (GOLDBERG, 1995, p. 74). Por isso, diz-se que a construção herdeira é um tipo mais específico da construção mãe.

Ao adotar a herança normal entre construções, Goldberg (1995) apresenta uma abordagem que subdivide esse processo em quatro, cada qual representado por um tipo de link, quais sejam, o link de Polissemia, o de Subparte, o de Instanciação e o de Extensão Metafórica. Presume-se que os links de herança, tais como as construções, apresentem estrutura interna – já que conectam partes específicas da construção mãe a partes específicas da filha – e sejam relacionados hierarquicamente – visto que a construção mãe domina a construção filha (1995, p.75).

Assim sendo, o link de Polissemia faz referência a construções que compartilham propriedades sintáticas, porém a contraparte semântica diverge. Em outras palavras, esse link captura relações entre um sentido particular de uma construção e as extensões que tal sentido pode adquirir.

O exemplo utilizado por Goldberg (1995) é o da construção *Ditransitiva*, cujo padrão sintático é associado a uma família de sentidos relacionados, os quais surgem a partir de um sentido central, representado pelo esquema “X CAUSA Y a RECEBER Z”, conforme a sentença (17). Dessa forma, outros diferentes esquemas de sentido surgem através do link de polissemia, tais como “X PRETENDE CAUSAR a Y RECEBER Z”, em (18); condições de satisfação implicam que “X CAUSA Y a RECEBER Z”, em (19); e “X CAUSA Y a NÃO RECEBER Z”, em (20).

- (17) Maria me deu a bola.
- (18) Maria me assou um bolo.
- (19) Maria me prometeu um bolo.
- (20) Maria me negou um pedacinho do bolo.

Por outro lado, para identificar relações em que uma construção é uma subparte de outra construção, são assinalados os links de Subparte. A construção [SN V_{Estar} SADJ para S_{inf}] é herdada pela construção de *Tempo_Iminente*, definida pelo esquema sintático [SN V_{Estar} para S_{inf}], por um link de Subparte, visto que, na construção herdeira, o SADJ não consta. A sentença (21) ilustra a construção herdada e (22), a de *Tempo_Iminente*. Note-se que a perda do constituinte gera alteração de significado.

- (21) Maria está pronta pra começar a faculdade.
- (22) Maria está pra começar a faculdade.

Por sua vez, os links de Instanciação demarcam situações em que uma construção é uma versão mais especificada de outra (GOLDBERG, 1995, p.79). É o que ocorre com a construção *Aspectual_Inceptiva* preenchida por “romper”, conforme definida por Sigiliano (2013). Comparando-se os exemplos em (23) e (24), nota-se que o preenchimento do *slot* de auxiliar pelo verbo “romper” traz à construção um sentido específico de início abrupto, o que não se verifica no caso em que a posição de auxiliar é preenchida por “começar”.

- (23) Maria começou a chorar.
- (24) Maria rompeu a chorar.

Finalmente, há o link de Extensão Metafórica, responsável por sinalizar quando construções interagem através de mapeamento metafórico. Exemplo desse link se encontra na análise de Salomão (2008) para a construção *Modal_Impessoal_com_Dar*, que emerge, motivada por uma complexa rede de heranças, como extensão metafórica da construção de *Transferência_de_Posse*, conforme se explica, a seguir:

A motivação crítica nesse processo é a conceptualização metafórica da alteração sociodinâmica, introduzida pela Transferência da Propriedade de um Recurso, como Habilitação virtual de um Agonista para a consecução de uma Finalidade. Outras metáforas ontológicas (Eventos

como Objetos; Causas como Aquisição/Perda de Propriedades; Finalidades como Destinações) contribuem para que situações progressivamente mais abstratas venham a ser evocadas nesses termos (SALOMÃO, 2008, p.108).

Logo, observem-se os exemplos (25) e (26), que representam, respectivamente, as construções de *Transferência_de_Posse* e a *Modal_Impessoal_com_Dar*, a qual se apresenta como extensão metafórica da primeira.

- (25) O Antônio deu o livro dele prá mim.
- (26) Dá pra ele chegar a tempo.

Contudo, Kay (2005, p.2) afirma que uma abordagem monotônica, isto é, baseada em unificação, é capaz de abranger um maior número de fatos semânticos e sintáticos por meio de um aparte teórico mais econômico. Observemos a análise realizada pelo pesquisador acerca do aparato teórico utilizado por Goldberg (1995).

Conforme se viu acima, um link de Polissemia é estabelecido para abarcar, por exemplo, a relação entre a construção *Ditransitiva* e uma de suas extensões de sentido, também conhecida como construção *Benefactiva*. Enquanto a primeira construção apresenta sentido central traduzido pelo esquema “X CAUSA Y A RECEBER Z”, na segunda, a causação é pretendida, mas não necessariamente realizada. Desse modo, a construção *Benefactiva* é vinculada à *Ditransitiva* por um link de Polissemia, visto que herda o esquema sintático central, ao passo que o sentido é uma extensão do significado original, segundo a análise de Goldberg (1995).

Kay (2005, p.4) aponta que, ao todo, Goldberg estabelece seis diferentes sentidos para a construção *Ditransitiva* – os quais se baseiam no sentido central –, de modo que cada sentido se combina com verbos de uma classe semântica distinta. Como exemplo, o sentido central combina-se com verbos de “dar” (tais como *give*, *hand* e *pass*), com verbos de causação instantânea de movimento balístico (*throw* e *toss*), e com verbos de causação contínua em uma direção especificada deiticamente (*bring* e *take*). Kay argumenta que postular vários sentidos da construção *Ditransitiva* e, ao mesmo tempo, reconhecer

classes semânticas de verbos é redundante no que se refere à contabilização das diferenças de implicação.

Em oposição, Kay (2005, p.4) propõe três subconstruções, de maneira que boa parte das distinções em implicações virá da semântica dos verbos apenas. Visto que a construção que licencia o destinatário como argumento sintático nuclear de verbos que não o exigem é uma construção de estrutura argumental, ela será chamada de *Construção de Destinatário* (*RC – Recipient Construction*). Assim, sentenças da *RC* produzem conjuntos diferentes de implicações dependendo da classe semântica do verbo.

As sentenças de (27) a (32) (KAY, 2005, p.2) exibem as três *RC* máximas propostas por Kay: a *RC Pretendida* (27), a *RC Direta* (28) e a *RC Modal* (29) a (32).

(27) A famous sculptor carved my sister a soap statue of Bugs Bunny.

(28) The catcher threw Pat the bean bag.

(29) The boss promised me a raise.

(30) The administration always denies late arrivals permission to enter.

(31) Aunt Maude bequeathed me a collection of risqué postcards.

(32) The referee allowed Kim two free throws.

(33) *My sister was carved a soap statue of Bugs Bunny by a famous sculptor.

Dois fatos motivam a distinção entre a *RC Pretendida* e as outras duas, segundo a explicação de Kay (2005, p.6): primeiramente, devido ao fato de apenas nesta construção o argumento destinatário não poder ser realizado como sujeito da passiva – conforme ilustra a sentença (33) –; também, o destinatário almejado precisa se beneficiar do recebimento do tema. Já a *RC Direta* se distingue das demais ao agregar a implicação de que o ator causa, de modo intencional, que o tema se mova e que o receptor putativo realmente o receba. Por fim, a *RC Modal* não apresenta restrições que limitem passivas, mas requerem a semântica beneficiária da *RC Pretendida*, além de compartilhar com ela a falta de uma implicação de recebimento.

Conforme Kay (2005, p.7) esclarece, na *RC Modal*, o ato realizado pelo ator incorpora uma intenção que envolve o recebimento do tema, mas em cada

subcaso, o evento de recebimento está sujeito a uma qualificação. Por exemplo, em verbos como *guarantee* e *promise*, o evento de recebimento está sujeito a uma obrigação do ator. Já em verbos como *refuse* ou *deny*, o evento de recebimento está sujeito à negação. Nesse sentido, Kay (2005, p.7) assevera que as diferenças de implicação entre os sentidos propostos por Goldberg (1995) podem ser interpretadas de acordo com os significados dos verbos, sem que seja necessário estabelecer mais variações da construção. Em suma, “cada verbo unificado com a *RC Modal* fornece sua própria modalização particular do evento de recebimento” (KAY, 2005, p.7) ²⁹.

Dessa forma, Kay (2005, p.19-20) conclui que vários sentidos postulados na análise de Goldberg são supérfluos, bem como os links não monotônicos estabelecidos entre as diferentes construções de destinatários. Reforça, ainda, que a abordagem monotônica é suficiente para explicar as diferenças de implicação, sendo desnecessário o conceito de polissemia. Assim, ao tentar se afastar do lexicalismo gerativista, Goldberg acaba por criar um aparato teórico redundante, desconsiderando que muitas diferenças por ela sinalizadas decorrem, na verdade, da semântica dos próprios verbos, e não de uma variação na construção.

No Constructicon, o modo completo (monotônico) é o empregado na estruturação da herança entre construções, conforme se demonstrará no capítulo 5, uma vez que se trata de um modelo computacional e, neste caso, trabalhar com a possibilidade de incluir exceções e regularidades parciais acarretaria uma série de indeterminações que poderiam, inclusive, inviabilizar a elaboração de um sistema de redes de herança.

Há, ainda, uma distinção entre cópia virtual e cópia real de informação. Na cópia real, Goldberg (1995, p.74) explica que as construções dominadas contêm toda a informação presente nas construções dominantes, de modo que cada construção é totalmente especificada, mas é redundante na medida em que a informação é herdada de construções dominantes. Em contrapartida, na cópia virtual, construções dominadas são especificadas apenas parcialmente de sorte que a informação herdada é armazenada somente na construção dominante. Assim, para se determinar por completo as características de uma

²⁹ Each verb unified with the Modal RC furnishes its own particular modalization of the reception event.

construção, é necessário buscar na árvore de herança as propriedades das construções dominantes. O Constructicon trabalha, na modelagem do recurso, com a cópia virtual, já que este modo é mais econômico e menos redundante, do ponto de vista de armazenamento de dados.

As relações de herança entre construções sinalizam, portanto, as conexões possíveis entre os elementos que compõem a rede de construções modelada. Entretanto, para além de modelar as possibilidades, faz-se necessário modelar também as restrições que operam sobre as construções. De modo relacionado, no contexto da aplicação do recurso construcional, além de ser necessário definir os casos licenciados por uma dada construção, é fundamental diferenciar aqueles que não o são. Nesse contexto, os conceitos de “precisão” e “recall” contribuem para essa separação, os quais são aplicados da seguinte forma:

Para cada rótulo que você está usando para identificar elementos nos dados, o conjunto de dados é dividido em dois subconjuntos: um que é rotulado como "relevante" para o rótulo, e um que não é relevante. A *precisão* é uma métrica que é calculada como a fração das instâncias corretas dentre aquelas que o algoritmo rotulou como estando no subconjunto relevante. O *recall* é calculado como a fração de itens corretos entre aqueles que realmente pertencem ao subconjunto relevante (PUSTEJOVSKY & STUBBS, 2013).³⁰

Em outras palavras, a precisão consegue atribuir X a todos os membros da categoria X, ao passo que o recall não inclui outros elementos nessa categoria. Ambas as métricas, ou seja, precisão e recall, são importantes para a adequação descritiva de um recurso. Nesse sentido, os modelos de construções, bem como as relações que estabelecem com outras construções, precisam ser propostos de modo a dar conta do conjunto de construtos licenciados pela construção e somente deles, evitando-se que uma dada construção modelada possa gerar supergeneralizações

Uma das abordagens para a identificação dos fatores que inibem as supergeneralizações por parte dos falantes atribui papel proeminente ao conceito de entrincheiramento. Croft & Cruse (2004, p.292) definem que será

³⁰ For each label you are using to identify elements in the data, the dataset is divided into two subsets: one that is labeled “relevant” to the label, and one that is not relevant. Precision is a metric that is computed as the fraction of the correct instances from those that the algorithm labeled as being in the relevant subset. Recall is computed as the fraction of correct items among those that actually belong to the relevant subset.

descrita como entrincheirada uma forma de palavra que ocorra com frequência suficiente para ser armazenada independentemente. Por isso, a hipótese que defendem é a de que “o armazenamento de uma forma de palavra, regular ou irregular, é uma função de sua frequência token” (CROFT & CRUSE, 2004, p.293)³¹. Dessa forma, exemplificam que mesmo uma instância regular do esquema de plural de substantivos como “*boys*” pode ser entrincheirada, visto que possui uma elevada frequência token. Por outro lado, o grau de entrincheiramento (produtividade) de esquemas – representações de regras tais como a marcação em –s para plural –, é determinado pela frequência type.

Diversos teóricos argumentam que o processo de entrincheiramento desempenha função essencial para restringir supergeneralizações, visto que crianças tendem a supergeneralizar itens usados com pouca frequência. Contudo, Goldberg (2006, p.94) alega que os efeitos imputados a tal fenômeno seriam mais bem atribuídos à preempção estatística, juntamente ao papel do contraste semântico ou pragmático.

A pesquisadora (GOLDBERG, 2006, p.94) afirma que uma maneira de supergeneralizações serem minimizadas baseia-se na ideia de que itens mais específicos são produzidos preferencialmente em vez de itens que são licenciados mas são representados de forma mais abstrata, desde que os itens compartilhem as mesmas restrições semânticas e pragmáticas.

Embora tal ideia seja familiar no âmbito morfológico, os casos de preempção entre duas formas sintagmáticas demandam uma análise mais aprofundada, já que expressões formadas a partir de construções sintagmáticas distintas nunca são idênticas semântica e pragmaticamente (GOLDBERG, 2006, p.95). Goldberg (2006, p.96), então, sugere as sentenças (34) e (35) para uma reflexão: por que a paráfrase preposicional da construção *Ditransitiva* em (35) impede o uso de (34), que é mal formada?

(34) *She explained me the story.

(35) She explained the story to me.

A teórica argumenta que uma forma estatística de preempção poderia desempenhar importante papel ao se aprender a evitar expressões como (34):

³¹ ... the storage of a word form, regular or irregular, is a function of its token frequency.

em uma situação em que se pode ter esperado que a construção A fosse enunciada, o aprendiz pode inferir que a construção A não é apropriada se a construção B é ouvida de modo consistente (GOLDBERG, 2006, p.96). Em outros termos, se o falante quisesse usar a outra formulação A, ele o teria feito. Logo, uma vez que ele não o fez, talvez tenha usado a formulação alternativa por uma razão.

Em termos de um Constructicon, modelar as restrições que se aplicam às construções é, portanto, tão necessário quanto modelar as possibilidades de relações entre elas. Conforme se verá no capítulo 5, tais restrições podem envolver uma série de aspectos, desde o material lexical que deve ser usado em uma construção, até as restrições semânticas que operam a partir dele ou sobre a construção como um todo. É justamente sobre a relação entre construções e frames que se debruça a próxima seção.

3.2 Relações entre Frames e Construções

Em seu texto de 1977, Fillmore rebate críticas à sua proposta de abordagem semântica a partir da teoria de casos e, ao fazê-lo, estabelece em sua argumentação a gênese do que mais tarde se tornaria a FrameNet. É nesse contexto que autor cunha a expressão “significados são relativizados a cenas”, lema da semântica de frames.

Fillmore (1977, p.60) pontua que os casos profundos concernem o eixo sintagmático, em vez do paradigmático, uma vez que compõem relações semânticas entre os elementos da sentença estabelecidas em um contexto, ao invés de se tratar de um sistema de oposições que diferenciam constituintes no eixo paradigmático.

Ressalta-se que esta teoria estava comprometida em determinar relações gramaticais nucleares em uma sentença, a saber, o sujeito, o objeto e o objeto indireto, de modo que oferecia descrições da valência semântica de verbos e adjetivos (FILLMORE, 1977, p.60). Mais tarde, esse seria o cerne da análise de sentenças no contexto da FrameNet, já que a inserção de camadas compreende – conforme visto na subseção 2.3.1 – a definição das funções gramaticais encontradas, bem como os elementos de frame presentes – entre outras propriedades.

Destacava-se o frame de caso, cuja função era a de vincular as descrições de situações a representações sintáticas, o que era feito através da atribuição de papéis semântico-sintáticos a determinados participantes da cena representada pela sentença (FILLMORE, 1977, p. 61). Dessa forma, Fillmore traçou propostas acerca do nível de organização de uma oração, relevantes para o significado e estrutura gramatical, fato que proporcionou uma maneira de descrever aspectos da estrutura lexical (FILLMORE, 1977, p. 62).

Outro importante fundamento apresentado por Fillmore nesse artigo é a ideia de perspectiva. Além dos três tipos de funções de constituintes da sentença elencados por Katz (1972, p.113), quais sejam, a gramatical – a qual inclui noções como sujeito e objeto –, a retórica – que abarca oposições como tópico e comentário – e a semântica – que estabelece elementos como agente e recipiente –, Fillmore alegou haver uma quarta maneira de visualizar tais funções por meio da noção de perspectiva (1977, p.60-61). Para ele, havia partes de uma mensagem que estavam em perspectiva enquanto outras estavam fora de perspectiva – noção fundamental para a teoria de casos e, mais tarde, também para a Semântica de Frames e para a Gramática de Construções.

Retomando o conceito de relatividade dos significados conforme a cena em questão, Fillmore (1977, p.72) assinalou que o número de casos não corresponde ao conjunto de noções necessárias para se analisar um estado ou evento, como se verifica ao se considerar a cena do evento comercial. Nela, há dois indivíduos envolvidos de forma agentiva: o comprador e o vendedor. Logo, não se pode prever a quantidade e quais casos existem independentemente do contexto da sentença, da cena evocada. Entretanto, ao se falar sobre tal evento, será exigido do falante que ele assuma uma determinada perspectiva e, de acordo com essa, certas entidades serão mencionadas, bem como certos verbos serão escolhidos – por exemplo, se for tomada a perspectiva do vendedor e da mercadoria, o verbo “vender” será empregado, conforme exemplifica Fillmore.

Nesse contexto, Fillmore traça uma definição inicial para o conceito de evocação de um frame:

O estudo da semântica é o estudo das cenas cognitivas que são criadas ou ativadas por enunciados. Sempre que um falante usa qualquer um dos verbos relacionados ao evento comercial, por exemplo, a cena inteira do evento comercial é colocada em jogo – é "ativada" – mas a

palavra particular escolhida impõe a esta cena uma perspectiva particular (FILLMORE, 1977, p. 73).³²

Em seguida, o linguista explicita o que sua célebre frase, “significados são relativizados a cenas”, denota. Ele explica que escolhemos e entendemos expressões ativando em nossas mentes cenas dentro das quais a palavra ou expressão tem uma função de nomeação ou descrição ou classificação.

É no mesmo texto que serão encontradas as primeiras bases da teoria que mais tarde seria chamada de Gramática das Construções. Todavia, embora essa e a Semântica de Frames sejam teorias irmãs, poucos trabalhos em Gramática das Construções traçam uma abordagem que sistematiza formalmente o conceito de frame. Há, obviamente, a exceção dos trabalhos guiados pelo modelo da BCG (KAY & FILLMORE, 1999) e daqueles desenvolvidos por pesquisadores ligados a uma *framnet* (MICHAELIS & RUPPENHOFER, 2001; BOAS, 2003; HASEGAWA ET AL., 2010). Isso não quer dizer que as demais abordagens neguem que os significados das construções sejam propostos em termos de frames (ou esquemas), mas, salvo algumas exceções, elas não usam um repositório estruturado e hierárquico de frames. Uma das exceções é a tentativa de Goldberg (2010), como se verá a seguir.

Goldberg (2010) discute a relação entre frames, verbos e construções ao argumentar que a única restrição na combinação de eventos designados por um único verbo é a de que tais eventos constituam um frame semântico coerente e estabelecido. Assim, ela demonstra que um verbo pode indicar subeventos não vinculados por uma relação de causa (assim como o verbo *blanch*³³) e pode, ainda, especificar modo e resultado simultaneamente (como *schuss*³⁴), fatos que haviam sido contestados por alguns estudiosos – tais como Croft (1991) e Rappaport-Hovav & Levin (2010). Contudo, a combinação de verbo e construção pode designar uma predicação única que não corresponda a um frame estabelecido.

³²The study of semantics is the study of the cognitive scenes that are created or activated by utterances. Whenever a speaker uses any of the verbs related to the commercial event, for example, the entire scene of the commercial event is brought into play – Is "activated" – but the particular word chosen imposes on this scene a particular perspective.

³³ O verbo *blanch* refere-se à imersão de alimentos em água fervente e depois em água fria, para remover a pele ou aumentar a cor (GOLDBERG, 2010, p.6).

³⁴ O verbo *schuss* significa esquiador direto para baixo (mudança de localização/resultado) intencionalmente e muito rápido (modo) (GOLDBERG, 2010, p.9).

Nesse contexto, Goldberg traça a definição do que seja um frame semântico de predicação, a saber, “um estado ou evento generalizado, possivelmente complexo, que constitui uma unidade cultural. Certos aspectos do frame semântico são perfilados; o resto constitui o pano de fundo” (2010, p.2) ³⁵.

A autora afirma que a única restrição acerca de eventos a serem rotulados por um verbo é a do frame convencional:

Para que uma situação seja rotulada por um verbo, a situação ou a experiência podem ser hipotéticas ou históricas e não precisam ser diretamente experienciadas, mas é necessário que a situação ou experiência evoquem uma unidade cultural familiar e relevante para aqueles que usam a palavra (GOLDBERG, 2010, p.11) ³⁶.

Entretanto, a partir de sua análise, a pesquisadora argumenta que os significados mais específicos que surgem a partir da combinação de uma construção de estrutura argumental e um verbo específico não exigem um frame semântico estabelecido ³⁷. Logo, verbos podem se combinar com construções de um modo inteiramente novo.

Assim sendo, pode-se perceber como o conceito de frame é empregado por Goldberg de maneira informal, uma vez que não são contemplados em sua definição os elementos que compõem o frame e o distinguem de outras situações. Essa mesma tendência se observa na maioria dos trabalhos em Gramática das Construções, conforme pontuado por Boas (2014).

Por outro lado, implementações computacionais da Gramática das Construções têm olhado para essa relação – e também para as relações de herança – de modo um pouco mais sistemático, conforme se verá na sessão seguinte, apesar de nenhum deles utilizar, como se propõe nesta tese, uma FrameNet como proposta de modelo semântico para dar conta do aporte semântico das construções.

3.3 Implementações Computacionais

³⁵ “(...) a generalized, possibly complex state or event that constitutes a cultural unit. Certain aspects of the semantic frame are profiled; the rest constitutes the background frame”.

³⁶ “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.”

³⁷ Exemplifica tal constatação a sentença “*I microwaved her some left over noodles*”, que denota “pretender dar algo preparado por microondas”. Seria implausível afirmar que tal significado pertence a um frame estabelecido (Ver GOLDBERG, 2010).

Nesta seção, apresentamos três implementações computacionais que, de alguma forma, abordam as relações entre frames e construções: o ECG Analyzer, o Constructicon de Berkeley e o Constructicon do Sueco. Essas três implementações foram escolhidas pelos motivos que se apresentam a seguir:

- i O ECG Analyzer é um parser com objetivos semelhantes ao do CARMA, sendo desenvolvido pela FrameNet Brasil e, apesar de não trabalhar propriamente com frames, implementa a noção de esquemas pareados a construções.
- i O Constructicon de Berkeley foi o projeto piloto a partir do qual derivam todos os demais constructica.
- ii O Constructicon do Sueco é o único desses, além do da FrameNet Brasil, que se encontra de fato implementado.

3.3.1 O ECG Analyzer

O ECG Analyzer é a implementação computacional da Embodied Construction Grammar (BRYANT, 2008). Trata-se de um analisador sintático-semântico que decompõe sentenças em termos de construções e esquemas utilizados. Baseia-se numa gramática computacional modelada segundo os princípios da ECG abordados resumidamente no capítulo 1.

No que tange às relações entre os polos formal e semântico, a ECG dispõe de quatro maneiras de sinalizar relações entre suas estruturas: por meio de papéis, de subtipagem, de restrições e de evocação de uma estrutura, como aponta Bryant (2008, p.35).

Os papéis nomeiam partes de uma estrutura, como se nota na Figura 22. Nela, observa-se, na primeira caixa à esquerda, o esquema TL, cujos papéis são *trajectory* e *landmark*. Tendo em vista que o esquema – ou estrutura – é análogo ao frame, os papéis são, portanto, comparáveis ao EFs. Há que se ressaltar, entretanto, que os esquemas da ECG são menos linguisticamente ancorados do que os frames em uma framenet. Esquemas incluem qualquer tipo de sistema conceitual, inclusive os muito abstratos, tais como Dentro-Fora, Orientação Vertical e Escala. Já os frames, em sua forma de implementação presente nas

framenets, tendem a ser mais linguisticamente ancorados, sendo que frames não lexicais cumprem, na rede, um papel mais organizacional e hierarquizante.³⁸

A subtipagem, por sua vez, permite a herança. Assim, nota-se, na Figura 22, que o esquema SPG é um subtipo – ou subcaso – do TL, assim como o esquema Motion é um subtipo do Process.

Já as restrições de tipo restringem os preenchimentos dos papéis, e o operador *self* permite que um esquema ou uma construção se refira a si próprio. No caso do esquema Motion, a restrição demonstra que, neste subcaso específico de Process, o papel *protagonist* do esquema de Process precisa ser um *mover*, conforme se verifica na Figura 22.

O autor lembra que evocar uma estrutura não significa que o elemento evocador estabelece com a estrutura evocada uma relação de parte-de ou de subtipo. Tal fato é aclarado por meio do conceito de hipotenusa, o qual só faz sentido em referência a um triângulo retângulo, embora a hipotenusa não seja um tipo de triângulo retângulo, nem o triângulo retângulo seja uma parte da hipotenusa. Logo, a relação de evocação é utilizada a fim de indicar o vínculo entre a hipotenusa e seu triângulo retângulo (BRYANT, 2008, p.35-36). Assim, como assinala a Figura 22, a relação de evocação marca, no esquema MotionAlongAPath, que o esquema SPG compõe seu significado, na medida em que o *mover* tem como restrição ser um *trajector*, nos termos do esquema SPG, ou seja, um *trajector* cujo *landmark* será especificado como um *path*.

Visto isso, a Figura 22 exhibe um subconjunto simplificado de esquemas da ECG para representar o significado da sentença “*Eve walked into the house*”. Bryant explica que o esquema Motion Along A Path (movimento ao longo de um caminho) usa a relação de evocação (Evokes) para ativar uma instância do esquema SPG (*Source Path Goal* – Fonte Caminho Alvo), com o intuito de representar um caminho e, em seguida, relacionar o papel do *mover* (aquele que move) ao papel do spg.trajector, o que significa que o *mover* é o *trajector* movendo-se ao longo de um caminho (2008, p. 36).

³⁸ Como se verá no capítulo 5, a partir das contribuições desta tese, na FrameNet Brasil, frames não lexicais são também recrutados por construções de estrutura argumental.

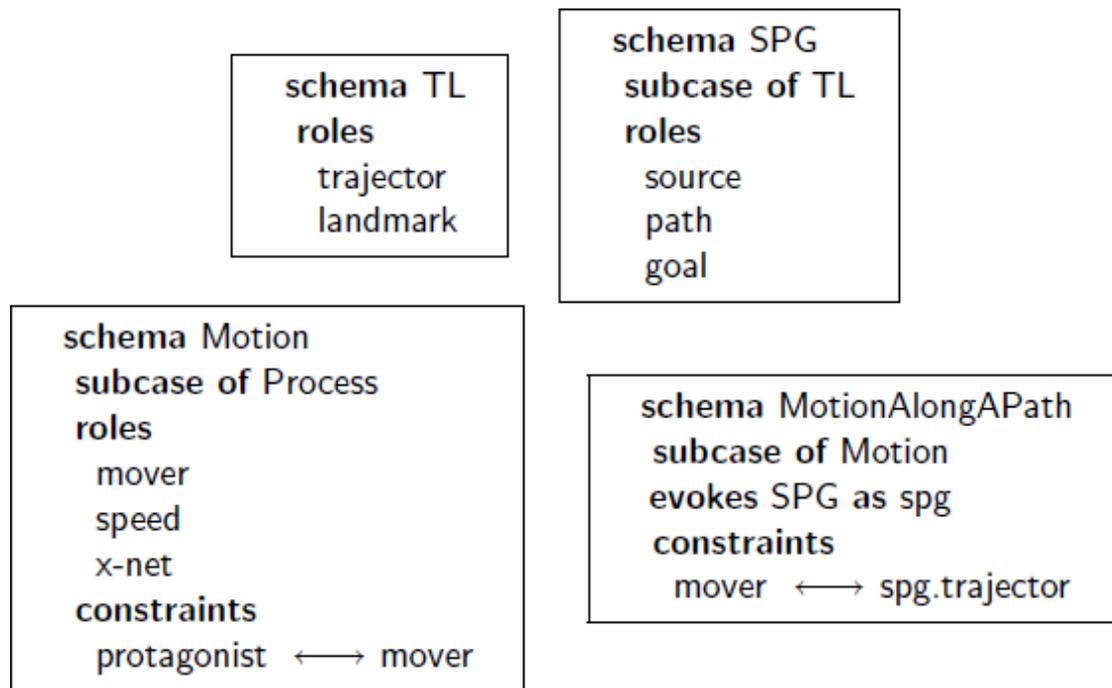


Figura 22: Subconjunto simplificado de esquemas da ECG para representar o significado da sentença “*Eve walked into the house*” (BRYANT, 2008, p.35).

Além dessas relações, Bryant (2008, p. 36) aponta que cada construção é definida a partir de três blocos, que podem ou não constar na definição: um bloco construcional, um da forma e um do significado. O bloco construcional especifica a constituição, os traços construcionais e as restrições. O bloco da forma, por sua vez, estabelece as restrições de ordem nos constituintes. Por fim, o bloco do significado especifica a denotação que a construção apresenta e adiciona qualquer significado não contemplado até então.

Dentro do bloco da forma, dois operadores atuam para restringir o ordenamento relativo dos constituintes: o *before*, que indica que o argumento à esquerda seja anterior ao que está à direita – podendo haver algum componente que se posicione entre eles –, e o *meets*, que exige que o argumento à esquerda venha imediatamente antes daquele que está à direita (BRYANT, 2008, p. 37). Tais restrições encontram equivalentes diretos no Constructicon da FrameNet Brasil, conforme se mostrará adiante.

Já o bloco do significado faz referência à cena que a construção em questão denota. No caso da construção Motion Along A Path – conforme se observa na Figura 23 –, a cena homônima é quem compõe o significado da

construção, e o significado de cada um dos constituintes desta está vinculado ao papel correspondente na cena (BRYANT, 2008, p. 37).

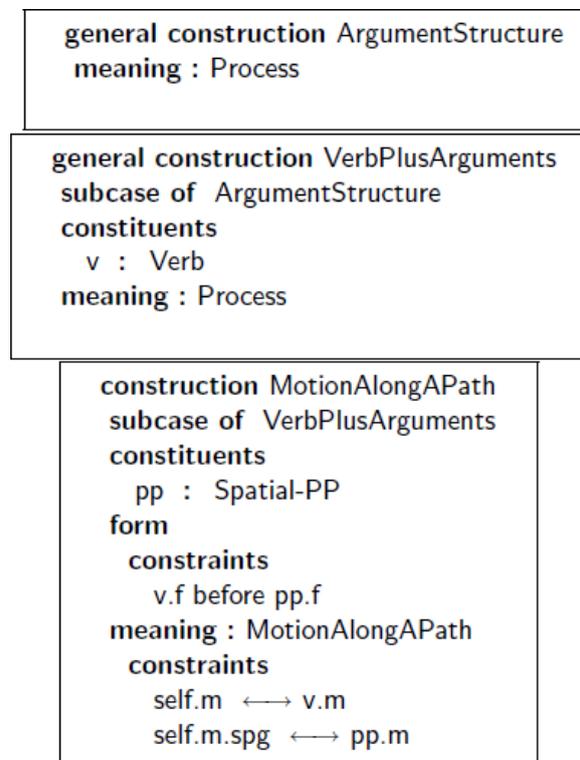


Figura 23: Construção MotionAlongAPath(BRYANT, 2008, p.38).

Como se verá no capítulo 5, esta tese implementa algumas relações semelhantes às do ECG Analyzer, como, por exemplo, a de Meaning, chamada por nós de Evocação, e a Subcase_of, chamada de Herança.

3.3.2 O Constructicon de Berkeley

O Constructicon de Berkeley, iniciado pelo projeto Beyond the Core, tinha por intuito abarcar casos não contemplados pela proposta de análise lexical da FrameNet – conforme assinalado no capítulo 2. Embora esses, portanto, sejam projetos irmãos, Lee-Goldman & Petruck (no prelo) ressaltam que eles ainda não estão totalmente integrados: é preciso representar as conexões explícitas entre frames e construções. Alguns passos, contudo, já começaram a ser sinalizados.

Nesse sentido, propõe-se um link entre construções e frames de modo que o significado de uma construção é especificado como sendo o significado de

um frame específico. Logo, como explicam Lee-Goldman & Petruck (no prelo), o significado da construção *Rate*, como em “*He types 75 words per minute*” é o significado do frame de *Rate*. Esse link, no entanto, ainda não foi estabelecido no recurso de Berkeley, de modo a conectar as duas bases de dados, da FrameNet e do Constructicon.

O vínculo entre construções e frames também não havia sido implementado no âmbito da FrameNet Brasil, anteriormente a esta tese, mas as limitações geradas pela ausência de tal link revelaram a urgência em estabelecê-lo. Lage (2013, p.105) discute os problemas gerados pela não interação entre os recursos, exemplificando com a construção *Modal_Epistêmica*, presente na sentença (36):

(36) Dá para dormir na viagem?

A construção que licencia (36) evoca o frame de *Possibilidade*, descrito na FrameNet. Entretanto, sem se estabelecer o link, ainda que haja menção, na descrição da construção, à evocação desse frame, os dados não estarão conectados. Considerando o fato de que o Constructicon do PB se dirige, sobretudo, a tarefas em tecnologia da linguagem, servindo inclusive, como já mencionado, de banco de dados para o CARMA, a não vinculação dos dados implica em uma perda de informações quando do processamento dessas. Como o CARMA interpretará sentenças, oferecendo boas análises, se não for reconhecido que o padrão construcional em (36) reflete, na verdade, um frame não previsto pela combinação das partes que compõem a construção?

Em conformidade, Lee-Goldman & Petruck (no prelo) confirmam que estabelecer analogias entre o Lexicon e o Constructicon da FrameNet parece natural e necessário. De tal modo, a Figura 24 expõe possíveis analogias entre os dois recursos:

FrameNet Lexicon	FrameNet Constructicon
Frame	Construction
Frame Evoking Element (FEE) (Lexical Unit, LU)	Construction Evoking Element (CEE)
Frame Element (FE)	Construction Element (CE)
lexicographic annotation	constructicographic annotation

Figura 24: Análogos Lexicon-Constructicon (LEE-GOLDMAN & PETRUCK, no prelo)

Lee-Goldman & Petruck (no prelo) explicam que a UL, ou EEF (Elemento Evocador do Frame), aponta o frame relevante para a interpretação da sentença em questão, do mesmo modo que o EEC indica a construção que licencia a estrutura linguística sob análise. EFs e ECs são análogos na medida em que são elementos distintivos do frame e da construção, respectivamente. Frames são comparáveis a construções, visto que são as categorias em torno das quais cada recurso se organiza, havendo, inclusive, possibilidade de estruturação por herança. Contudo, também se pode comparar a construção à UL, já que ambas são capazes de evocar frames. De acordo com os pesquisadores,

esses análogos surgem naturalmente quando se consideram os tipos de informação que ULs, construções e frames codificam, juntamente com as conexões entre eles. Esse desenvolvimento natural apenas enfatiza a necessidade de desenvolver frames e construções em paralelo e os benefícios de estabelecer conexões profundas entre o lexicon e o constructicon (LEE-GOLDMAN & PETRUCK, no prelo)³⁹.

Relacionado a tais aspectos, está o propósito de integrar todo o repertório de construções do inglês, transformando-o em uma rede. Para tanto, alguns links foram propostos, no domínio especulativo. A herança, por exemplo, relacionará uma construção filha à construção mãe, caso todas as restrições formais, semânticas e pragmáticas da mãe também sejam verdadeiras para a filha – possibilitando adicionar restrições à filha. O Constructicon de Berkeley também

³⁹ These analogues arise naturally when considering the types of information that LUs, constructions, and frames encode, along with the connections between them. This natural development only emphasizes the need to develop frames and constructions in parallel and the benefits of establishing thorough connections between the lexicon and constructicon.

propõe herança entre ECs – o EC filho herda do EC pai, nos mesmos moldes da herança entre construções. Ressalta-se que essa relação sempre faz parte da herança entre construções. Ainda, propõe-se o estabelecimento de um link entre EC e construção, quando um EC pertencente a uma construção instancia outra (LEE-GOLDMAN & PETRUCK, no prelo), link este já modelado pela FrameNet Brasil, conforme se demonstrou na seção final do capítulo 2.

Assim sendo, embora o Constructicon de Berkeley reconheça a necessidade de se representar as relações entre frames e construções, isso não foi implementado, de maneira que o recurso continua desvinculado do lexicon. Lage (2013) argumenta que a inexistência de relações entre construções cria lacunas no recurso. É o que ocorre com a construção do inglês *Wherewithal*, cuja descrição sinaliza que essa herda da construção *Relative_clause*. Entretanto, tal construção nem mesmo existe no Constructicon do inglês, pois ainda não foi descrita no sistema. Dessa forma, a ausência dos links de herança impede que o Constructicon de Berkeley, da forma como hoje se apresenta, seja definido como uma rede, tal como a FrameNet.

3.2.3 O Constructicon do Sueco

Outra importante implementação computacional é o Constructicon sueco (SweCcn), o qual, além de ser projetado como ferramenta linguística e como recurso para tecnologia de linguagem, também se dirige ao ensino de sueco como língua estrangeira. Conforme explicam Lyngfelt et al. (no prelo), esse recurso é uma combinação entre uma versão prática – simplificada – da Gramática das Construções e inspirações metodológicas advindas da lexicografia, tanto da produção de dicionários como da própria FrameNet. Por isso, a entrada de uma construção deve ser formalizada, concisa e de fácil manuseio.

Nesse contexto, cada entrada pode conter até quinze campos de informação, os quais se subdividem em três grupos, sendo o primeiro aquele que apresenta uma descrição geral da construção a partir destes cinco campos: nome, ilustração, definição, esboço da estrutura e exemplos. Primeiramente, o nome, o qual identifica a construção, deve ser significativo para um leitor humano, e a ilustração traz um exemplo idealizado da construção em questão.

O campo da definição é preenchido em prosa aos moldes de descrições de dicionário, apontando o significado da construção. Por sua vez, o esboço da estrutura exibe uma representação simples da composição gramatical da construção, que pode incluir classe de palavra, tipo sintagmático, função gramatical e item lexical. Por fim, os exemplos – instâncias autênticas retiradas de *corpora* – são disponibilizados para ilustrar o modo como a construção é usada em termos de variabilidade, de estrutura sintática e de informação (LYNGFELT ET AL., no prelo).

Um segundo grupo oferece uma categorização das construções bem como uma caracterização dos elementos da construção, por meio dos campos: categoria, tipo, elementos da construção, palavras-chave e palavras comuns. No primeiro campo, é atribuída a toda construção uma categoria básica conforme o tipo sintagmático da construção como um todo, e, no segundo, cada uma é classificada em tipos baseados em uma propriedade saliente, tais como: propriedades funcionais (contraste, comparação), estrutura geral (coordenação), elementos da construção específicos (reflexivo, elementos implícitos) e foco do aluno (para construções consideradas particularmente relevantes para estudantes de sueco). No campo seguinte, descrevem-se os elementos da construção quanto às propriedades formais e semânticas. As palavras-chave correspondem aos Elementos Evocadores da Construção, e, em palavras comuns, expõe-se uma lista de unidades lexicais comumente usadas na construção (LYNGFELT ET AL., no prelo).

O terceiro grupo faz menção à organização interna do SweCcn, bem como a referências externas, através destes campos: herança, frame, ID de Berkeley, comentário e referência. Em herança, aponta-se a relação com outras construções no SweCcn, ao passo que, em frame, refere-se ao frame evocado pela construção – quando essa o faz, já que esse recurso construcional reconhece casos de construções que não evocam frames. O ID de Berkeley identifica uma entrada correspondente no Constructicon de Berkeley. Finalmente, o campo comentário sinaliza observações sobre a análise, e referência indica trabalhos para análises (LYNGFELT ET AL., no prelo).

Como o SweCcn foi desenvolvido independentemente da FrameNet sueca, há algumas consequências em relação aos padrões incluídos nesses recursos. Na FrameNet Brasil, por exemplo, um algoritmo explícito é empregado

para distinguir quais padrões linguísticos devem ser analisados na FrameNet, e quais são destinados ao Constructicon – conforme visto na subseção 2.3.2. Visto que no SweCcn as construções são selecionadas sem se considerar a FrameNet sueca, há uma sobreposição parcial de dados entre os dois recursos (LYNGFELT ET AL., no prelo). Além disso, os frames apontados no campo relevante das construções suecas são apenas links para o relatório de visualização do frame na base de dados da Swedish FrameNet++. Nenhum tipo de conexão estruturada (entre ECs e EFs, por exemplo) existe nesse recurso.

Embora ainda não tenha sido implementada no SweCcn uma rede que conecte as construções já descritas, Lyngfelt et al. (no prelo) sinalizam esse como um objetivo central, tanto para favorecer a estrutura interna do banco de dados, quanto por razões teóricas. Para tanto, apontam a necessidade de descrever construções mais gerais como nós centrais na rede, e de estabelecer relações interconstrucionais, verticalmente, por meio da herança, e horizontalmente, por meio de links de subparte responsáveis por relacionar ECs que partilham traços.

Consideradas as discussões traçadas nesse capítulo, notar-se-á que a implementação de relações e do editor de restrições no Constructicon da FrameNet Brasil contribui significativamente para a evolução desse tipo de recurso, no sentido de que preenche importantes lacunas deixadas pelos projetos congêneres, na medida em que:

- i. propõe-se fazer uso de uma base de frames já desenvolvida e atestada em termos lexicais, ao invés de criar novas representações para cada um dos aspectos semânticos e lexicais das construções. Nesse sentido, avança em relação ao ECG Analyzer, que precisa criar novos esquemas e novas construções para cada elemento na base.
- ii. propõe-se modelar, no banco de dados, as relações de herança entre construções e sua constituência. Nesse sentido, avança em relação ao Constructicon de Berkeley que, apesar de reconhecer a necessidade de tais relações, não as implementa.
- iii. propõe-se estabelecer links modelados entre as bases de dados de frames, unidades lexicais e construções. Nesse sentido, avança em relação aos constructicons de Berkeley e do Sueco, visto que

ambos apenas sinalizam frames e ULs que aparecem nas construções de modo indicativo.

Nas seções que se seguem, demonstra-se como esses avanços foram alcançados no âmbito da pesquisa objeto desta tese.

4 METODOLOGIA

A metodologia empregada neste trabalho aplica a proposta de Dias-da-Silva et al. (2007) ao domínio da FrameNet Brasil; nesse sentido, divide-se nas fases Linguística, Representacional e Implementacional. Anteriormente à pesquisa relatada nesta tese, o Constructicon da FrameNet Brasil não possuía relações que interligassem as construções em uma rede, tampouco era conectado à base de dados lexical. É válido ressaltar que o trabalho aqui exposto foi desenvolvido ao longo de cinco anos e, portanto, devido à demanda de que essas relações fossem postas em práticas no recurso construcional, isso foi sendo implementado no decorrer desses anos de pesquisa. Tal fato permitiu que outros trabalhos concernentes ao Constructicon, como o de Almeida (2016), testassem, através do CARMA, as relações modeladas no âmbito desta pesquisa de tese.

Nesse contexto, a análise de construções a serem tratadas pelo Constructicon conduziu esta pesquisa ao levantamento das relações e restrições que deveriam ser incluídas nesse recurso. Como é tendência na FrameNet Brasil, construções já estudadas por diversos linguistas são escolhidas para serem modeladas no Constructicon. Assim, a construção de *Argumento_Cindido*, definida por Sampaio (2010), foi modelada no trabalho de Almeida (2016), juntamente com as construções *Transitiva_Direta_Ativa* e *Ergativa*, ao passo que Marção, Torrent & Matos (2017) têm se dedicado à modelagem das construções *Interrogativas-QU*.⁴⁰

Entretanto, no que concerne ao emprego das relações de Herança Construcional e Evocação, bem como às restrições baseadas em frames, todos esses modelos confiam nas propostas oriundas desta tese, as quais, por sua vez, tiveram como ponto de partida, na **Fase Linguística**, a análise das construções da família Para Infinitivo, estudada por Torrent (2009; 2015), a já mencionada *Dativo_com_Infinitivo*, a qual foi desenvolvida por Laviola (2015), bem como sobre a construção *Aspectual_Inceptiva* determinada por Sigiliano (2013).

⁴⁰ Estas e todas as demais construções modeladas no Constructicon da FrameNet Brasil podem ser consultadas em <http://www.ufjf.br/framenetbr/dados/constructicon/>

A construção de *Dativo_com_Infinitivo* torna ainda mais clara a necessidade de se transformar o Constructicon em uma rede, uma vez que essa construção se relaciona com outras da mesma família – e, também, externas à família – por redes de herança construcional, as quais capturam mais claramente informações herdadas e compartilhadas.

Para sustentar a rede de herança, foi necessário instituir construções em nível abstrato, que não se realizam concretamente. Por essa razão, ao se estabelecer uma construção é preciso sinalizar se essa é abstrata ou não, como visto na seção anterior.

Além da rede de heranças, era uma necessidade vincular a base de dados construcional à lexical, com o intuito prático de proporcionar análises mais completas para tarefas em NLU, mas também com a finalidade de unir teorias irmãs.

Por sua vez, a construção *Aspectual_Inceptiva* (SIGILIANO, 2012, 2013) apresenta distintas propriedades que conduziram ao estabelecimento de novas restrições – além daquelas já definidas por Almeida (2016), quais sejam, CE_meets, CE_before, inspiradas pelo ECG Analyzer (BRYANT, 2008), e CE_construction – a fim de que o reconhecimento dessa por máquinas pudesse ser efetuado de modo bem-sucedido.

Levantadas as análises linguísticas e os problemas que impõem à modelagem, esta pesquisa se debruçou sobre a **Fase Representacional**, a qual consiste em, dadas as estruturas de dados do Lexicon, do Constructicon e da rede de frames, estabelecer de que modo as propriedades das construções descritas deveriam ser armazenadas na base dados.

Uma vez definidos os requisitos, passa-se à **Fase Implementacional**, em que a pesquisadora se reúne com a equipe de desenvolvimento de software da FrameNet Brasil e solicita a adição da relação e/ou restrição na FrameNet Brasil WebTool, a ferramenta de gestão de banco de dados e anotação desenvolvida pelo projeto. Depois de implementadas, testam-se as adições e, em seguida, elas são aperfeiçoadas a fim de acomodar ajustes necessários.

Na sequência, modelam-se as construções no Constructicon e quando relevante e/ou possível, desenham-se experimentos – como o de Almeida (2016) – que possam testar o modelo através do CARMA.

A Figura 25 resume a metodologia desta tese.

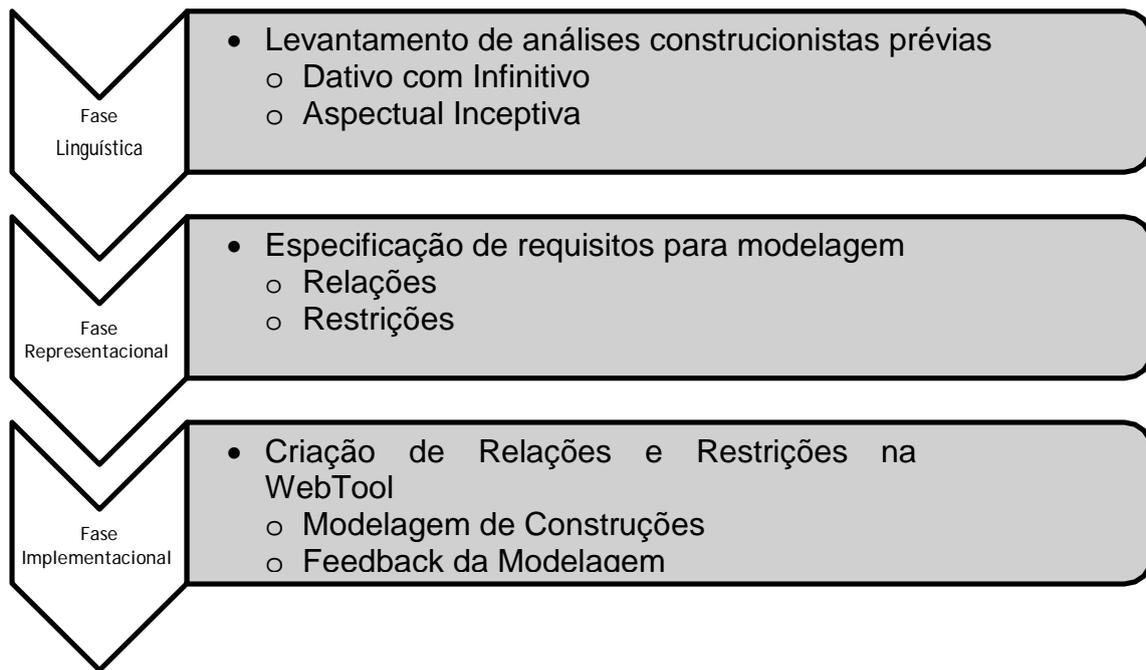


Figura 25: Processo de criação de novas relações e restrições no Constructicon da FrameNet Brasil

5 MODELAGEM DAS RELAÇÕES ENTRE CONSTRUÇÕES E ENTRE FRAMES E CONSTRUÇÕES NO CONSTRUCTICON DO PORTUGUÊS BRASILEIRO

Este capítulo tem por objetivo apresentar a modelagem de relações e restrições no Constructicon da FrameNet Brasil a fim de que um tratamento adequado das construções a tarefas de NLU seja possibilitado. Na primeira seção, será discutida a transformação do Constructicon em uma rede através da relação de Herança construcional. A segunda seção mostrará a união das bases de dados da FrameNet Brasil e do Constructicon, estabelecida por meio da relação de Evocação e pela instituição de novas restrições no Editor de Restrições.

5.1 Relações entre Construções

Esta seção se dedica à explicitação das discussões e decisões tomadas para a escolha da estrutura de herança levada a cabo no Constructicon da FrameNet Brasil, bem como à demonstração da modelagem dessa relação em tal recurso.

Retomando as três etapas de desenvolvimento que envolvem o processo de formação de um sistema de PLN sugeridas por Dias-da-Silva et al. (2007), a Fase Linguística já foi desempenhada previamente a este trabalho por Torrent (2009; 2015) e Laviola (2015) e será apresentada na subseção 5.1.2, quando a construção tomada para a modelagem da relação de herança for descrita.

Portanto, a contribuição desta tese se concentra nas Fases Representacional e Implementacional. A Fase Representacional será demonstrada na subseção 5.1.1, onde será exposta a elaboração conceitual do sistema de herança, o qual servirá para representar os resultados encontrados na fase anterior, a Linguística. Já a Fase Implementacional, a qual também será exibida na subseção 5.1.2, é encaminhada para ser modelada pela equipe de programação de que dispõe a FrameNet Brasil. Nessa fase, exibe-se a integração conceitual e de estrutura de dados dos componentes envolvidos.

5.1.1 Herança Construcional

No capítulo 3, foram apresentadas propostas para lidar com a Herança entre construções. Conforme afirmado, o Constructicon da FrameNet Brasil adotará, a partir desta tese, a herança múltipla, da mesma forma que todas as abordagens em Gramática das Construções. Como se observa, as construções interagem de maneira a herdar de mais de uma construção, agregando propriedades de cada uma que consta em sua linhagem.

O Constructicon também trabalhará com o modo completo de herança. A discussão do capítulo 3 evidenciou que a proposição dos links não monotônicos (modo normal) de Goldberg (1995) é desnecessária e redundante. Além disso, como visto, devido às aplicações computacionais a que se destina o Constructicon da FrameNet Brasil, o modo normal poderia ocasionar a inviabilização de um sistema de herança.

Vejamos a seguinte situação hipotética: se for estabelecido que a construção B herda de A, mas que há traços de A que não estão presentes em B, ou que estes podem ser ignorados por uma razão X, o sistema passa a não ter parâmetro, podendo realizar conexões indevidas. Como uma das principais finalidades do Constructicon é servir de base de dados para alimentar um analisador sintático-semântico, o CARMA, essa espécie de imprecisão é extremamente indesejável, sobretudo porque, para realizar análises sintático-semânticas, são necessárias restrições. As construções de *Dativo_com_Infinitivo*, por exemplo, herdam, simultaneamente, todos os traços constantes da *Oração_relativa* e da *Para_Infinitivo*, de acordo com o modo completo de herança.

Ademais, como o CARMA trabalha com ativação propagada – Spreading activation, metodologia matemática que modela a propagação de energia por uma rede neural – (MATOS ET AL., 2017), se um link de herança é proposto, isso significa que um caminho conectando dois nós na rede foi aberto para que a carga de energia, que é representada por um valor numérico, possa passar de um nó para o outro. Se trabalhamos com herança completa, sabemos que todos os caminhos abertos entre dois nós por herança serão do mesmo tipo, logo, podemos atribuir a eles todos um mesmo peso para reduzir ou reforçar a carga energética durante a passagem, subtraindo ou somando algum valor pré-

definido ao número que representa a quantidade de energia. Porém, se a herança é normal, cada um desses links pode ser muito diferente de cada um dos outros, de modo que não é possível modelá-los como sendo de um só tipo.

Por fim, vale confirmar que a cópia virtual é adotada na estruturação da herança no Constructicon, de maneira que a informação herdada é armazenada apenas na construção dominante, evitando, assim, a redundância no armazenamento de dados. Porém, ao importar essas informações para o CARMA passa-se a trabalhar com a cópia real de informações, para poupar gasto de energia ao sistema.

Assim se constitui a relação de herança que estruturará a rede construcional. Observar os caminhos percorridos na composição de cada construção é fundamental, pois confere ao Constructicon a possibilidade de armazenar informações de forma econômica e não redundante. Pode-se, por exemplo, inferir conhecimentos acerca de uma construção, sabendo-se de quais outras construções ela herda. Na próxima seção, será demonstrado de que maneira os princípios aqui adotados são modelados no Constructicon, a partir da já referida construção de *Dativo_com_Infinitivo*.

5.1.2 Implementação de Exemplo: Construções Relativas

Uma das construções que inspirou e serviu de base de análise para instituição das relações entre construções foi a *Dativo_com_Infinitivo*. Para compreendê-la melhor, é preciso antes considerar a família à qual tal construção pertence, que, por sinal, foi a primeira testada quando da implementação do Constructicon (LAGE, 2013). Como tal família apresentava grande variedade no que diz respeito à tipologia construcional, exibindo desde construções de estrutura argumental aberta até aquelas lexicalmente especificadas, a Para Infinitivo se mostrou ideal para iniciar o recurso construcional, já que diversas possibilidades poderiam ser testadas. Assim, a família Para Infinitivo contribuiu para o estabelecimento dos critérios definidores dos materiais que deveriam ou não ser analisados no Constructicon.

Esse conjunto de construções, definido por Torrent (2009, 2011, 2015), compartilha propriedades sintático-semânticas e, por essa razão, compõe uma família. No tangente à contraparte sintática, essas construções apresentam o

esquema [SN₁ V SA/SN₂ para (SN₃) V_{INF}] ou se relacionam a ele por meio de diferentes links herança. Observe-se que toda a família herda propriedades da construção *Adjuntiva_de_Finalidade*, cujo esquema formal é [para (SN) V_{INF}]. Quanto ao aspecto semântico, tais construções evocam o frame de *Finalidade*, ou algum frame a ele relacionado, visto que o V_{INF} representa um objetivo a ser alcançado. Vale observar, conforme assinala Torrent (2015), que o frame de *Finalidade* também se relaciona à preposição “para”, classificada tradicionalmente na gramática como responsável por introduzir orações infinitivas de finalidade.

Torrent (2009, 2015) explica que a família Para Infinitivo se relaciona em uma rede de disposição radial, em que se localiza o protótipo, conectado aos outros membros da rede através de links de herança. Tais links são propostos segundo a herança de modo normal, conforme Goldberg (1995). Nessa rede, as construções são organizadas em uma hierarquia de esquematicidade-lexicalidade, de modo que as construções mais esquemáticas são definidas em relação ao sentido geral da rede e à sintaxe mais esquemática; as construções de classe específica são determinadas quando um sentido mais específico é combinado a uma classe particular de verbo ou adjetivo que preencha o slot do verbo finito ou do sintagma adjetivo, respectivamente; e as construções especificadas lexicalmente ocorrem somente se verbos ou adjetivos específicos preenchem o slot de verbo finito ou do sintagma adjetivo.

Qualificada como a construção mais esquemática da família, Torrent (2015) esclarece que a *Dativo_com_Infinitivo* herda a estrutura da construção de *Transferência_de_Posse* por um link de Extensão Metafórica, conforme a análise de Salomão (1990, 2007) a respeito da construção de *Habilitação* com o verbo “dar”. Ao mesmo tempo, se combina com a *Adjuntiva_de_Finalidade*, cujo esquema [para (SN) V_{INF}] preenche o espaço do argumento oblíquo da *Transferência_de_Posse*, de forma que o SN dativo concentra os papéis de Beneficiário da transferência e Agente do V_{INF}.

Desse modo, a construção de *Dativo_com_Infinitivo* evoca o frame de *Suficiência*, segundo o qual um recurso habilita um beneficiário/agente a atingir uma finalidade. Diferentes maneiras de se

disponibilizar esse recurso serão codificadas em conformidade com a semântica do V_{FIN} empregado, que pode variar consideravelmente: há verbos de transferência, como “dar” e “passar”; de atividade mental, como “conhecer” e “ver”; e de existência, como “ter” e “existir”. As sentenças de (37) a (42) exemplificam tais casos.

- (37) A Vânia **deu** dinheiro pra mim comprar comida.
- (38) Maucha **passou** o número pra gente ligar.
- (39) A Natália **conhece** alguma simpatia para dormir?
- (40) Adrieli **viu** uns hotéis pra ficarmos.
- (41) Eles **têm** lugar pra ficar?
- (42) **Existem** recursos para financiar o projeto?

Uma análise mais aprofundada (LAVIOLA & TORRENT, 2012; TORRENT, 2015) da construção de *Dativo_com_Infinitivo* em comparação com a construção *Wherewithal* do inglês apontou novos achados. Segundo Fillmore (2008), nesta construção, o SN_2 é um recurso suficiente para a execução da atividade indicada pelo V_{INF} . A sentença (37), anteriormente mencionada, evidencia tal afirmativa: o SN “dinheiro” é o recurso suficiente para que o beneficiário/agente “mim” alcance a finalidade de “comprar comida”, codificada pelo V_{INF} . Assim, corrobora-se a afirmação de que a *Dativo_com_Infinitivo* evoca o frame de Suficiência.

Contudo, Laviola (2015) pontua que nem todas as construções de *Dativo_com_Infinitivo* oferecem interpretação habilitativa. Uma análise do *corpus* de Torrent (2009) mostrou que, de 127 sentenças contendo esse padrão construcional, 24 apresentavam uma interpretação deôntica. Neste grupo, embora a estrutura sintática seja muito similar à de *Dativo_com_Infinitivo*, o sentido diverge, uma vez que há obrigação de se realizar uma atividade. A sentença (43) evidencia tais casos, já que o SN_2 “provas”, associado ao V_{INF} , aponta para uma ação a ser realizada obrigatoriamente.

- (43) A professora tem provas pra corrigir.

Laviola (2015, p.85) destaca, ainda, que não poderia ser argumentado que a interpretação divergente decorre da complementação entre o V_{INF} e o núcleo nominal do SN_2 , visto que a sentença (44) também dispõe dessa

complementaridade, mas ambas as leituras são possíveis, tanto a deôntica quanto a habilitativa.

(44) Maria tem dinheiro para gastar.

Nesse contexto – e após a verificação de que se tratava da mesma construção separada apenas pela diferença na evocação do frame –, a construção de *Dativo_com_Infinitivo* foi dividida em duas, a *Dativo_com_Infinitivo_Suficiência* e a *Dativo_com_Infinitivo_Obrigaçã*o. Assim, o Constructicon da FrameNet Brasil propôs definições para cada construção, conforme as Figuras 26 e 27 demonstram.

Dativo_com_infinitivo_Suficiência [Sufficiency_NP]

Definição	
Uma Para_Sinf indica um propósito cuja realização depende de um recurso codificado em um NP_Núcleo . A valência do NP_Núcleo é aumentada de modo a requerer a Para_Sinf em uma estrutura semelhante à de relativização.	
Elementos da Construção	
NP_Núcleo [Head_NP]	Nome que codifica um recurso empregado para a realização do propósito indicado pela Para_Sinf .
Para_Sinf [Para_Sinf]	Oração que indica o propósito cuja realização é habilitada pelo recurso codificado pelo NP_Núcleo .

Figura 26: *Dativo_com_Infinitivo_Suficiência*

Dativo_com_infinitivo_Obrigaçã [Obligation_NP]

Definição	
Uma Para_Sinf indica uma obrigação cuja realização envolve um item codificado em um NP_Núcleo . O item indicado no NP_Núcleo deve, necessariamente, instanciar um EF nuclear que se mapeie a uma função temática não-agentiva no frame evocado pelo verbo infinitivo.	
Elementos da Construção	
NP_Núcleo [Head_NP]	Nome que codifica um item sobre o qual a ação codificada pelo verbo da Para_Sinf deve ser realizada.
Para_Sinf [Para_Sinf]	Oração que indica a obrigação que deve ser realizada sobre o NP_Núcleo .

Figura 27: *Dativo_com_Infinitivo_Obrigaçã*

Note-se que *DCI_Suficiência* e *DCI_Obrigaçã* possuem os mesmos ECs, *NP_NÚCLEO* e *PARA_SINF*; a diferença está na relação que se estabelece entre eles: na primeira construção, o *NP_NÚCLEO* é o recurso suficiente que habilita a realização da ação codificada na *PARA_SINF*, enquanto na segunda construção, o *NP_NÚCLEO* codifica um item sobre o qual recai a ação imputada na *PARA_SINF*.

Uma vez que a diferenciação entre as construções provém da contraparte semântica que apresentam, ambas as versões de *Dativo_com_Infinitivo* herdam das mesmas construções. A relação de herança direta se estabelece com a construção de *Cláusula_relativa_infinitiva_para* cuja definição no Constructicon é evidenciada na Figura 28.

Cláusula_relativa_infinitiva_para [Para_infinitival_relative]

Definição
Um NP_Núcleo é relativizado por uma Para_Sinf uma sentença infinitiva regida pela preposição "para".
Elementos da Construção
NP_Núcleo [Head_NP] O NP_Núcleo é o antecedente da Para_Sinf .
Para_Sinf [Para_Sinf] A Para_Sinf relativiza o NP_Núcleo .

Figura 28: *Cláusula_relativa_infinitiva_para*

A *Cláusula_relativa_infinitiva_para* também apresenta os mesmos ECs que as construções de *Dativo_com_Infinitivo*, porém a relação que se estabelece entre os ECs é apenas de relativização do *NP_NÚCLEO* pela *PARA_SINF*. Trata-se de uma construção abstrata, uma vez que foi proposta unicamente para capturar as propriedades compartilhadas por ambas as construções *DCI* e prover a ponte entre elas e duas outras construções de que herdam: a *Cláusula_relativa_infinitiva* e a *Adjuntiva_final_infinitiva*.

A *Cláusula_relativa_infinitiva*, por sua vez, difere da *Cláusula_relativa_infinitiva_para* pelo fato de a sentença infinitiva não ser regida pela preposição “para”, como mostra a definição na Figura 29. Ela serve de mãe, portanto, de toda e qualquer construção relativa reduzida de infinitivo no PB, regida por qualquer preposição. Assim, uma sentença como (45) é licenciada por *Cláusula_relativa_infinitiva_de*, que também herda de *Cláusula_relativa_infinitiva*.

(45) Maria tem panelas de ir ao forno.

Cláusula_relativa_infinitiva [cxn_relativa_infinitiva]

Definição
Um NP_Núcleo é relativizado por uma Sinf .
Elementos da Construção
NP_Núcleo [Head_NP] O NP_Núcleo é o antecedente relativizado na Sinf .
Sinf [Sinf] A Sinf relativiza o NP_Núcleo .

Figura 29: *Cláusula_relativa_infinitiva*

Finalmente, a *Cláusula_relativa_infinitiva* herda da *Cláusula_relativa*, na qual o *NP_NÚCLEO* é relativizado não mais por uma sentença infinitiva, mas por uma sentença relativa, o que é apontado na Figura 30 e exemplificado na sentença (46).

(46) Maria tem panelas que vão ao forno.

Cláusula_relativa [Relative_clause]

Definição
Um NP_Núcleo é relativizado por uma Srel .
Elementos da Construção
NP_Núcleo [Head_NP] O NP_Núcleo é o antecedente relativizado pela Srel .
Srel [Srel] A Srel relativiza o NP_Núcleo .

Figura 30: *Cláusula_relativa*

Conforme afirmado anteriormente, a *Cláusula_relativa_infinitiva_para* também herda da *Adjuntiva_final_infinitiva*, a qual compõe a sentença infinitiva regida pela preposição “para”. Esta construção se funde à *Cláusula_relativa_infinitiva* para constituir a *Cláusula_relativa_infinitiva_para*. A Figura 31 mostra a definição no Constructicon da *Adjuntiva_final_infinitiva*.

Adjuntiva_final_infinitiva []

Definição	
Cláusula hipotática final reduzida de infinitivo.	
Elementos da Construção	
Prep_para []	Preposição que rege a cláusula infinitiva.
Sinf [Sinf]	Cláusula infinitiva regida pela Prep_para. Evoca o EF Finalidade do frame de Finalidade.

Figura 31: *Adjuntiva_final_infinitiva*

Assim se organiza a estrutura de herança que sustenta a construção de *Dativo_com_Infinitivo_Suficiência* e *Dativo_com_Infinitivo_Obrigaçãõ*. Assim sendo, retomando a linhagem de construções tratadas nesta subseção, observa-se que, de fato, a *Cláusula_relativa_infinitiva_para* herda de mais de uma construção simultaneamente – da *Cláusula_relativa_infinitiva* e da *Adjuntiva_final_infinitiva* –, o que atesta a herança múltipla.

Verifica-se, também, o modo completo de herança, já que em todos os casos descritos a construção herda toda a informação constante da construção mãe, sem que haja exceções, de modo que cada construção filha é um modo mais específico da mãe, isto é, a *Dativo_com_Infinitivo_Suficiência* e a *Dativo_com_Infinitivo_Obrigaçãõ* são versões especificadas da construção de que herdam.

Nota-se, ainda, que a cópia virtual se estabelece. De tal modo, para se saber por completo algumas propriedades da *Dativo_com_Infinitivo_Suficiência*, por exemplo, é preciso investigar em sua hierarquia de herança as propriedades das construções de que herda.

Destarte, junto à descrição de cada construção no Constructicon, serão mencionadas as relações de herança em que se inserem tais construções, como mostram as Figuras 32 e 33.

Dativo_com_infinitivo_Suficiência [Sufficiency_NP]

Definição
Uma Para_Sinf indica um propósito cuja realização depende de um recurso codificado em um NP_Núcleo . A valência do NP_Núcleo é aumentada de modo a requerer a Para_Sinf em uma estrutura semelhante à de relativização.
Exemplo(s)
Elementos da Construção
NP_Núcleo [Head_NP] Nome que codifica um recurso empregado para a realização do propósito indicado pela Para_Sinf .
Para_Sinf [Para_Sinf] Oração que indica o propósito cuja realização é habilitada pelo recurso codificado pelo NP_Núcleo .
Relações
Herda de Cláusula_relativa_infinitiva_para

Figura 32: Herança na construção *Dativo_com_Infinitivo_Suficiência*

Ao visualizar a descrição da construção de *Dativo_com_Infinitivo_Suficiência* no Constructicon, o usuário visualizará a Figura 32. Após a exibição da definição, dos exemplos da construção e dos ECs, por fim, são expostas as relações desta construção, a qual herda apenas da *Cláusula_relativa_infinitiva_para*. Por outro lado, a Figura 33 mostra uma situação diversa.

Cláusula_relativa_infinitiva_para [Para_infinitival_relative]

Definição
Um NP_Núcleo é relativizado por uma Para_Sinf , uma sentença infinitiva regida pela preposição "para".
Exemplo(s)
Elementos da Construção
NP_Núcleo [Head_NP] O NP_Núcleo é o antecedente da Para_Sinf .
Para_Sinf [Para_Sinf] A Para_Sinf relativiza o NP_Núcleo .
Relações
Herdado por <i>Dativo_com_infinitivo_Obrigaçã</i> , <i>Dativo_com_infinitivo_Suficiência</i>
Herda de <i>Adjuntiva_final_infinitiva</i> , <i>Cláusula_relativa_infinitiva</i>

Figura 33: Herança na construção *Cláusula_relativa_infinitiva_para*

Em 33, o usuário notará que, além de herdar da *Adjuntiva_final_infinitiva* e da *Cláusula_relativa_infinitiva*, a *Cláusula_relativa_infinitiva_para* também é herdada pela *Dativo_com_Infinitivo_Suficiência* e

*Dativo_com_Infinitivo_Obrigaçã*o. Dessa maneira, é possível observar onde a construção em foco se posiciona na rede de herança.

Além da menção acerca das relações de herança que a construção estabelece, há ainda, no Constructicon, uma ferramenta conhecida como Grapher, através da qual é possível visualizar mais amplamente as relações de herança.

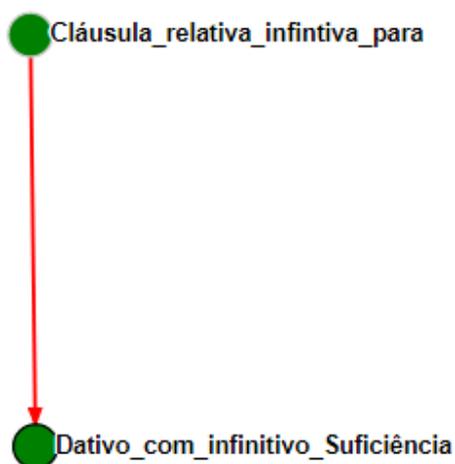


Figura 34: Grapher de *Dativo_com_Infinitivo_Suficiência*

Na Figura 34, observa-se que a *Dativo_com_Infinitivo_Suficiência* é a construção filha da *Cláusula_relativa_infinitiva_para*. Cada círculo verde indica uma construção, e a seta vermelha marca a relação de herança construcional. A construção para a qual a seta aponta é aquela que herda.

Na Figura 35, é possível notar como a herança se estabelece entre os ECs das duas construções. Os círculos brancos de bordas pretas indicam os ECs pertencentes a cada construção; as setas vermelhas atuam da mesma forma para indicar a herança, de maneira que o EC para o qual a seta aponta é o que herda.

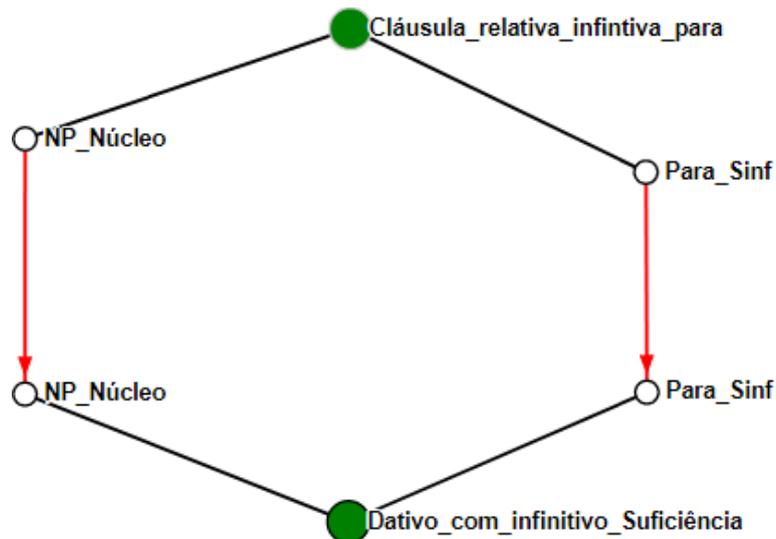


Figura 35: Grapher expandido de *Dativo_com_Infinitivo_Suficiência*

Do mesmo modo, verifica-se o Grapher com as relações de herança concernentes à *Cláusula_relativa_infinitiva_para*. Na Figura 36, visualizam-se as duas construções de que esta construção herda, a *Cláusula_relativa_infinitiva* e a *Adjuntiva_final_infinitiva*. Ademais, notam-se as duas construções herdeiras, a *Dativo_com_Infinitivo_Suficiência* e *Dativo_com_Infinitivo_Obrigaçã*.

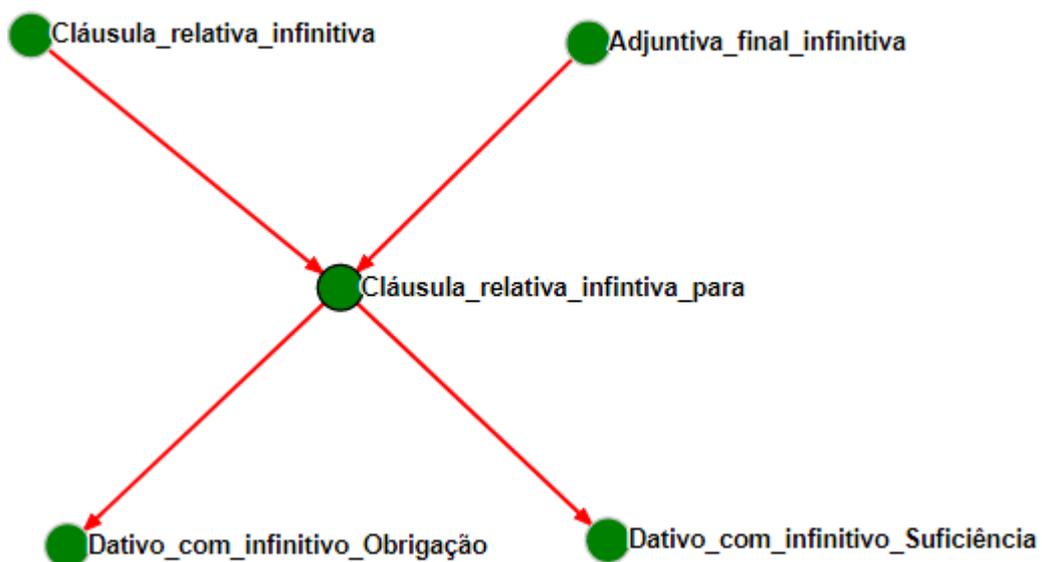


Figura 36: Grapher de *Cláusula_relativa_infinitiva_para*

Se o usuário desejar ver mais detalhes sobre as relações de herança, o Grapher expandirá revelando detalhes da relação. A Figura 37 exhibe os detalhes da herança entre a *Adjuntiva_final_infinitiva* e a *Cláusula_relativa_infinitiva_para*. Nessa figura, nota-se que o EC *PARA_SINF* herda dos dois ECs da *Adjuntiva*, quais sejam, a *SINF* e a *PREP_PARA*, ao passo que o EC *NP_NÚCLEO* não herda de nenhum EC da *Adjuntiva*, visto que ele provém da *Cláusula_relativa_infinitiva*.

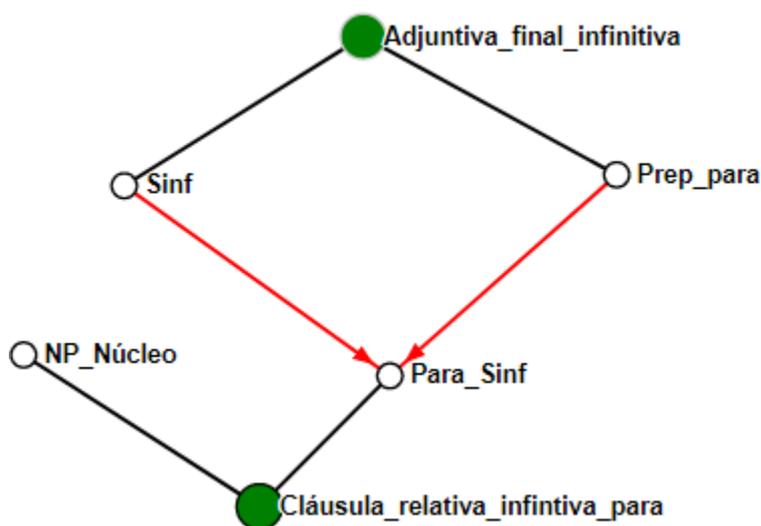


Figura 37: Grapher expandido de *Cláusula_relativa_infinitiva_para*

Por conseguinte, através da implementação da rede de herança entre construções, o Constructicon da FrameNet foi, afinal, convertido em uma rede. Essa transformação é um avanço frente a outros Constructicons já desenvolvidos para outros idiomas, uma vez que este é o primeiro recurso a integrar as construções de tal modo.

5.2 Relações entre Frames e Construções

Nesta seção, serão expostas as discussões que embasaram a estruturação da relação entre frames e construções no âmbito do Constructicon da FrameNet Brasil. Do mesmo modo, exhibe-se a proposta de modelagem dessa relação no referido recurso.

No que concerne à proposta que aqui se fará acerca da relação entre frames e construções, é válido destacar – lembrando a proposta de Dias-da-Silva et al. (2007) – que a Fase Linguística também já foi desenvolvida previamente a este trabalho quando Torrent (2009; 2015) e Laviola (2015) descrevem a construção de *Dativo_com_infinitivo_Suficiência* e *Dativo_com_infinitivo_Obrigaçã*o – já apresentada na subseção 5.1.2 – e quando Sigiliano (2012, 2013) realiza a definição da construção *Aspectual_Inceptiva* – a ser exibida na subseção 5.2.4.

Portanto, dedicamo-nos nesta tese, como mencionado, à Fase Representacional, a ser demonstrada nas subseções 5.2.1 e 5.2.3, quando discorreremos sobre a constituição conceitual da relação de Evocação e sobre as restrições semânticas de preenchimento de slot, respectivamente. A Fase Implementacional ficará a cargo das subseções 5.2.2 e 5.2.4, em que mostraremos a modelagem da relação e da restrição supra referidas.

5.2.1 Evocação

Um dos fatores motivadores para o desenvolvimento desta tese constituiu-se na pergunta de pesquisa apresentada na Introdução. Dessa forma, buscou-se verificar se o já estabelecido modelo semântico da FrameNet seria adequado à modelagem da contraparte semântica das construções descritas no Constructicon, uma vez que os itens lexicais são, também, um tipo de construção.

Usar o modelo de frames da FrameNet era fundamental porque, em primeiro lugar, esse é um recurso estabelecido e amplamente utilizado, com imenso repositório de frames já descritos. Logo, ao contrário do ECG Analyzer, que necessita criar, para cada construção nova a ser analisada, um novo esquema e descrevê-lo, a fim de arcar com a contraparte semântica, o Constructicon poderia dispor de um instrumento afim – já que a FrameNet é produto de uma teoria irmã – numa fase avançada de desenvolvimento. Dessa maneira, comparando-se, todo o trabalho gerado pelo ECG Analyzer poderia ser poupado no Constructicon, de modo a otimizar a produção de resultados.

Em segundo lugar, era fundamental relacionar os frames descritos na FrameNet às construções do Constructicon, pois, de tal modo, pela primeira vez,

frames e construções seriam modelados descritiva e computacionalmente para atuar juntos. O Constructicon incorporaria, portanto, a união sistemática de tais construtos, façanha há tanto tempo assinalada, mas nunca implementada em um recurso.

Nesse sentido, na busca por responder à mencionada pergunta de pesquisa, seguimos a hipótese de que a FrameNet se prestaria a modelar o aspecto semântico das construções. De tal modo, no capítulo 3, levantamos as propostas de implementações descritivas e computacionais que esboçam algum tratamento da relação entre frames e construções.

Nesse contexto, várias propriedades da ECG foram discutidas, inclusive a relação de Meaning, a qual serviu de inspiração para a criação da relação que uniria os dois construtos cognitivos, frame e construção. Embora a ECG não trabalhe efetivamente com frames, mas com esquemas, ela propõe uma relação que aponta o esquema acionado por uma construção. No Constructicon, essa relação ganhou o nome de Evocação, já que, na literatura em Semântica de Frames, materiais linguísticos evocam frames, e será responsável por relacionar a construção ao frame por ela evocado, nas situações em que isso ocorrer.

Assim, apontar especificamente o cenário evocado pela construção e a forma como ela compõe o significado através da relação de Evocação se mostra de grande relevância para o Constructicon da FrameNet Brasil. Diferentemente de outros Constructicons, como o sueco (LYNGFELT ET AL., no prelo), no qual não há problema se várias construções apresentarem o mesmo aporte semântico, já que não há o objetivo de se criar um parser fundamentado em seu recurso, o Constructicon brasileiro se subsume ao CARMA. Logo, supergeneralizações precisam ser evitadas a fim de que o modelo construcional licencie apenas construtos possíveis na língua e, conseqüentemente, o analisador nele fundado não produza resultados equivocados.

Diversamente da relação de herança entre construções, em que a construção que herda é filha, no sentido de receber todas as informações da construção mãe, a relação de Evocação não implica que o frame evocado seja pai da construção que o evoca, isto é, essa não deve conter toda a estrutura do frame.

5.2.2 Implementação de Exemplo: Construções de Dativo com Infinitivo

Como visto na seção anterior, as construções de *Dativo_com_infinitivo_Suficiência* e *Dativo_com_infinitivo_Obrigaçao* diferenciam-se em função da indicação do frame que evocam. Por sinal, a relação que se estabelece entre frame evocado e construção é outra que ainda não havia sido modelada no Constructicon e que se caracteriza como fundamental para a vinculação das bases de dados da FrameNet, com seu repertório de frames e unidades lexicais, e do Constructicon, com as construções.

Nesse contexto, observem-se as definições, dadas pela FrameNet Brasil, dos frames evocados pelas duas construções de *Dativo_com_infinitivo* – Figuras 38 e 39.

Suficiência

Definição	
Um Item está localizado em uma Escala relativamente a um valor crítico que é determinado por alguma Situação_habilitada . Geralmente, quando a Situação_habilitada não é expressa, ela é interpretada como o desejo de que o Item esteja localizado em um ponto particular da Escala . Com muitos dos adjetivos nesse frame, a Escala não é mencionada explicitamente. Nesses casos, a Escala é geralmente a quantidade incorporada no significado dos adjetivos.	
Elementos de Frame Nucleares	
Escala [Scale]	A Escala é o atributo do Item que está sendo descrito.
Item [Item]	A entidade cuja propriedade escalar é especificada.
Situação_habilitada [Enabled_situation]	Este EF identifica a Situação_habilitada , em relação à qual um valor crítico é julgado.

Figura 38: Frame de Suficiência

A Figura 38 apresenta a definição do frame de Suficiência, evocado pela *Dativo_com_infinitivo_Suficiência*, seguida da lista dos EFs, ESCALA, ITEM e SITUAÇÃO.

Ser_obrigado

Definição	
Sob alguma Condição , geralmente implícita, uma Parte_responsável deve realizar alguma Obrigaçã . Se a Obrigaçã não se cumpre, pode haver Consequências indesejáveis, que podem ou não ser ditas.	
Elementos de Frame Nucleares	
Obrigaçã [duty]	A ação que a Parte_responsável é obrigada a realizar. Eu tenho que tirar um visto para o Canadá.
Parte_responsável [responsible_party]	A pessoa que precisa realizar a Obrigaçã . Todos devem votar.

Figura 39: Frame de Ser_obrigação

Por sua vez, a Figura 39 exibe a definição do frame evocado pela *Dativo_com_infinitivo_Obrigaçã*, Ser_obrigado, acompanhado de seus EFs, OBRIGAÇÃO e PARTE_RESPONSÁVEL.

Assim sendo, a primeira modificação implementada a fim de estabelecer a relação de Evocação foi a sinalização, em cada construção, acerca do frame evocado, fato que se observa nas Figuras 40 e 41.

Dativo_com_infinitivo_Suficiência [Sufficiency_NP]

Definição	
Uma Para_Sinf indica um propósito cuja realização depende de um recurso codificado em um NP_Núcleo . A valência do NP_Núcleo é aumentada de modo a requerer a Para_Sinf em uma estrutura semelhante à de relativização.	
Elementos da Construção	
NP_Núcleo [Head_NP]	Nome que codifica um recurso empregado para a realização do propósito indicado pela Para_Sinf .
Para_Sinf [Para_Sinf]	Oração que indica o propósito cuja realização é habilitada pelo recurso codificado pelo NP_Núcleo .
Relações	
Evoca	Suficiência
Herda de	Cláusula_relativa_infintiva_para

Figura 40: Relação de Evocação na construção *Dativo_com_infinitivo_Suficiência*

Dativo_com_infinitivo_Obriga o [Obligation_NP]

Defini�o	
Uma Para_Sinf indica uma obriga�o cuja realiza�o envolve um item codificado em um NP_N�cleo . O item indicado no NP_N�cleo deve, necessariamente, instanciar um EF nuclear que se mapeie a uma fun�o tem�tica n�o-agentiva no frame evocado pelo verbo infinitivo.	
Elementos da Constru�o	
NP_N�cleo [Head_NP]	Nome que codifica um item sobre o qual a a�o codificada pelo verbo da Para_Sinf deve ser realizada.
Para_Sinf [Para_Sinf]	Ora�o que indica a obriga�o que deve ser realizada sobre o NP_N�cleo .
Rela�es	
Evoca	Ser_obrigado
Herda de	Cl�usula_relativa_infinitiva_para

Figura 41: Rela o de Evoca o na constru o *Dativo_com_infinitivo_Obriga o*

Nas Figuras 40 e 41, ap s a defini o e a exibiu o dos ECs, s o indicadas as rela es de que participa a constru o em foco. Al m da informa o a respeito dos links de heran a, consta, agora, o frame evocado.

Novamente, o Grapher ser  um instrumento  til para mostrar os detalhes da rela o de Evoca o. Desse modo, ao selecionar a constru o de *Dativo_com_infinitivo_Sufici ncia*, nota-se a rela o de heran a com a *Cl usula_relativa_infinitiva_para*, como mostrado na se o anterior, bem como o frame evocado, *Sufici ncia*. A rela o de Evoca o   indicada pela seta azul, que aponta para o frame evocado, sempre sinalizado pelo quadrado vermelho – Figura 42.

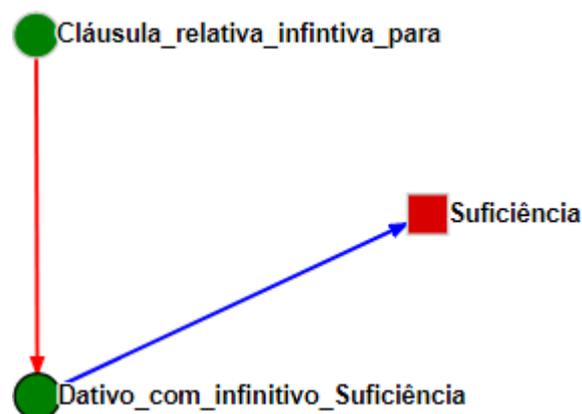


Figura 42: Grapher com a rela o de Evoca o na constru o *Dativo_com_infinitivo_Sufici ncia*

A relação de Evocação é criada no Editor de Relações da WebTool e permite ao analista, além de selecionar o frame evocado pela construção, dar detalhes sobre a estruturação dessa relação em termos de uma eventual correspondência entre ECs e EFs. A Figura 43 mostra como isso foi feito para a construção de *Dativo_com_infinitivo_Suficiência*.

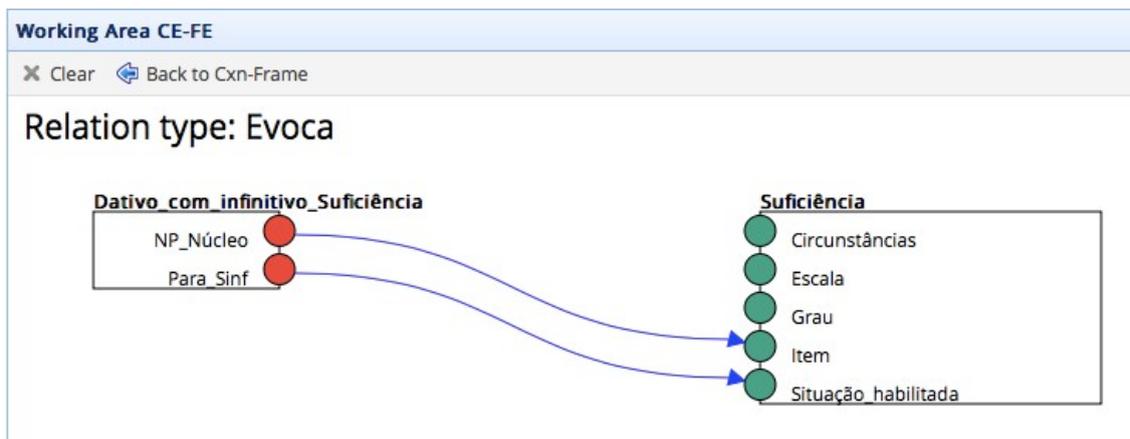


Figura 43: Editor de Relações e estruturação da relação de Evocação.

Uma vez definida a relação e sua estrutura, quando da anotação de construtos licenciados pela construção, a camada EF é preenchida automaticamente pelo sistema, a partir da escolha do anotador para a identificação dos ECs, conforme mostra a Figura 44.

		[13657] AST_MS_APP NI Ela deu dinheiro para mim viajar.									
CE											
Suficiência.FE											
CEE											
Other											
CstrPT											

Figura 44: Exemplo de tela de anotação construcional com atribuição automática de EFs.

Dessa forma, a relação implementada traz mais uma contribuição para a FrameNet Brasil, na medida em que semi-automatiza parte do processo de anotação construcional.

A relação de Evocação, entretanto, não é suficiente para dar conta de todos os aspectos semânticos das construções. Assim sendo, novas restrições foram adicionadas ao Editor, conforme se verá a seguir.

5.2.3 Restrição Semântica de Preenchimento de Slot

Mesmo depois de implementada a relação de Evocação, algumas propriedades ainda não eram capturadas, de maneira que a construção usada de teste não era plenamente descrita. De tal modo, além da relação de Herança entre construções e da relação de Evocação entre frames e construções, era preciso modelar restrições – além daquelas já estabelecidas no Constructicon – relativas à constituição da construção.

Em termos de um Constructicon, modelar as restrições que se aplicam às construções é tão necessário quanto modelar as possibilidades de relações entre elas. Tais restrições podem envolver uma série de aspectos, desde o material lexical que deve ser usado em uma construção, até as restrições semânticas que operam a partir dele ou sobre a construção como um todo.

Tais restrições buscam modelar tanto propriedades semânticas genéricas dos ECs, tais como permitir seu preenchimento apenas por itens lexicais de um determinado tipo semântico, quanto os efeitos da preempção estatística e/ou entrincheiramento, na medida em que modelam os aspectos únicos de uma dada construção que impedem que um parser tome construtos não licenciáveis por ela como o sendo.

Ressalta-se que as restrições semânticas de preenchimento de slot que serão apresentadas na próxima subseção não deixam de ser um tipo de relação entre frames e construções, estabelecendo conexão entre as bases de dados da FrameNet e do Constructicon. Entretanto, como operam diretamente sobre os ECs – e não sobre a construção como um todo –, nas discussões com a equipe de desenvolvimento de software, optou-se por incluí-las no Editor de Restrições, de modo a permitir ao analista incluir todas as propriedades dos ECs de uma única vez, ao modelar a construção.

5.2.4 Implementação de Exemplo: Construções Aspectuais Inceptivas

A construção *Aspectual_Inceptiva* (CI) (SIGILIANO, 2012) serviu de base para a implementação das restrições de preenchimento de slot discutidas na subseção anterior. Composta da estrutura sintática $[V1_{fin+} (prep) V2_{inf}]$, a CI apresenta uma série de peculiaridades, como veremos a seguir.

Nessa construção, a posição de $V1_{fin}$ pode ser ocupada por um grande número de verbos que não são tipicamente aspectuais, tais como “dancar”, “desatar”, “entrar” e “romper”. A respeito do $V2_{inf}$, como a linguista assinala (SIGILIANO, 2012, p. 131), em sua investigação, buscou-se elaborar uma lista de tipos semânticos nessa posição para observar suas semelhanças e regularidade de ocorrência. O objetivo desse levantamento era verificar os padrões de combinação entre $V1_{fin}$ e $V2_{inf}$ na *CI*. Foram encontrados 13 tipos semânticos⁴¹ na posição do $V2_{inf}$, conforme o Quadro 2:

Número	Tipo verbal	Exemplos
1	Manipulação	Ordenar, persuadir, fazer com que
2	Desejo	Querer, desejar, ter vontade de que
3	Percepção	Ver, ouvir, escutar, sentir
4	Processo mental	Saber, entender, compreender
5	Atitude proposicional	Achar, acreditar, crer
6	Elocução	Dizer, falar, contar
7	Deslocamento	Ir, vir, andar, entrar
8	Ação	Fazer, pegar
9	Estado / atributo	Ficar, ser, permanecer, estar
10	Fenômeno da natureza	Chover, nevar
11	Mudança de estado	Secar, passar [o tempo]
12	Sentimento	Amar, gostar, odiar
13	Emoção	Chorar, entristecer, soluçar [de tristeza].

Quadro 2: Tipos semânticos de $V2_{inf}$ (SIGILIANO, 2012, p.131)

Levantados os tipos semânticos encontrados em $V2_{inf}$, observou-se a ocorrência de cada um deles associada aos verbos que preenchem a posição de $V1_{fin}$, componente indicador do aspecto inceptivo. A Tabela 1 mostra essas associações em porcentagem e em número de dados.

⁴¹ Os tipos semânticos foram levantados com base em observações nos dados e em Noonan (2007) e Halliday (1994) *apud* Sigiliano (2012).

V2 _{inf}	V1 _{fin}				
	Danar a	Danar de	Romper a	Desatar a	Entrar a
Manipulação				0.42% (1)	
Desejo					0.6% (2)
Percepção					3.25% (11)
Processo mental		50% (1)		0.42% (1)	5.6% (19)
Atitude proposicional					2.07% (7)
Elocução			6.2% (4)	4.21% (10)	17.66% (60)
Deslocamento			7.8% (5)	9.7% (23)	7.37% (25)
Ação	66.6% (2)		18% (13)	9.7% (23)	53.4% (181)
Estado / Atributo					2.07% (7)
Fenômeno da natureza				1.3% (3)	0.31% (1)
Mudança de estado		50% (1)		0.8% (2)	
Sentimento					0.9% (3)
Emoção	33.3% (1)		68% (46)	73.41% (174)	6.78% (23)
TOTAL	100% (3)	100% (2)	100% (68)	100% (237)	100% (339)

Tabela 1: Distribuição das ocorrências da *CI* com base na combinação de tipos semânticos de V2_{inf} com cada possibilidade de preenchimento de V1_{fin} (SIGILIANO, 2012 p.132)

Sigiliano (2012, p. 133) destaca que as instanciações preenchidas em V1_{fin} por “entrar” se conectam, em sua maioria, a verbos de ação; por outro lado, aquelas que são preenchidas por “romper” e “desatar” se ligam, principalmente, a verbos de emoção. Assim, em alguns casos, as ocorrências sinalizam uma forte associação de um preenchimento específico de V1_{fin} a um tipo determinado de V2_{inf}, o que relembra a proposta de Goldberg (1995) de que os verbos não se inserem aleatoriamente nas construções.

A razão para o uso aspectual de verbos que não são tipicamente desse tipo provém do mecanismo de extensão⁴²:

Como a *CI*, ao longo do processo de gramaticalização, tem o significado de V1_{fin} esvaziado de seu conteúdo mais lexical e preenchido por um sentido gramatical de aspecto – fundado na semântica vestigial dos

⁴² De acordo com Heine & Kuteva (2005; 2006 *apud* SIGILIANO, 2012), mecanismo de extensão refere-se ao fato de que o uso em novos contextos motiva novos significados.

verbos que podem ocupar essa posição (...) –, a posição de $V2_{inf}$ passa a poder ser preenchida por verbos cuja combinação com os respectivos $V1_{fin}$ seria improvável, caso um verbo como entrar, por exemplo, mantivesse seu sentido lexical de ‘movimento para dentro’ (SIGILIANO, 2012).

Sigiliano aponta, ainda, para o esquema de representação do modelo de Quebra de Barreiras⁴³ como elemento que sustenta os usos aspectuais de $V1_{fin}$ e sinaliza que marcadores de aspecto associados à quebra repentina de barreira vinculam-se, principalmente, aos verbos de emoção. Sustenta-se que o processo de extensão da CI para abarcar $V2_{inf}$ do tipo emoção ocorre preferencialmente com os verbos “romper” e “desatar” devido à sua semântica vestigial da Quebra de Barreiras (2012, p. 140).

Ademais, duas metáforas de Lakoff & Johnson (2002 [1980] *apud* SIGILIANO, 2012) explicam a relação entre $V1_{fin}$ (“romper” e “desatar”) e $V2_{inf}$ (emoção), quais sejam:

EMOÇÃO É CONTÊINER: Estou num grande estado de nervos!

EMOÇÃO É LÍQUIDO NUM CONTÊINER: Ele transbordava de alegria.

Conforme explica Sigiliano, $V1_{fin}$ e $V2_{inf}$ têm associados em si a noção de contêiner, já que, “ao se romper algo, há uma barreira quebrada, e ao se desatar algo, também. Quando um contêiner é rompido ou desatado, as emoções ‘líquidas’ dentro dele tendem a se espalhar, deixando de ser contidas” (2012, p. 139).

Desse modo, Sigiliano conclui que os aspectuais “romper” e “desatar” devem combinar-se frequentemente com verbos de emoção e, no caso de ocorrer a extensão da CI para o uso com verbos de emoção, tais aspectuais devem figurar de modo prioritário nesses novos usos.

Para uma análise no âmbito do Constructicon, foi preciso considerar uma forma de modelar tal restrição. Para tanto, primeiramente, foram verificados os frames evocados pelos verbos da CI . Desse modo, podem-se enquadrar os tipos semânticos de $V2_{inf}$ como conjuntos de verbos que evocam um mesmo frame. Em outros termos, os verbos do tipo “emoção” evocarão um determinado frame, assim como os de “ação” evocarão outro frame, e assim por diante.

⁴³ Quebra de Barreiras é um dos padrões básicos da dinâmica de forças propostos por Talmy (2001, p. 415 *apud* SIGILIANO, 2012).

Como demonstra Sigiliano (2012), as restrições de coocorrência devem-se à persistência de traços semânticos – ou frames – nos verbos finitos. Assim, se observarmos os casos em que o $V1_{fin}$ é ocupado por “romper”, tal verbo descontextualizado evoca o frame de *Causar_fragmentar*, cuja definição se segue na Figura 45.

Causar_fragmentar

Definição
Um Agente , repentinamente e frequentemente violento, separa o Paciente_inteiro em dois ou mais Pedaços menores, resultando na não existência do Paciente_inteiro como tal. Muitos itens lexicais são marcados com um tipo semântico negativo, indicando que a fragmentação é necessariamente julgada como prejudicial ao Paciente_inteiro original. Compare este frame com Danificar, Tornar_não-funcional e Remover

Figura 45: Frame de *Causar_fragmentar*

O traço semântico capturado pelo frame de *Causar_fragmentar* permanece em “romper” e, através da metáfora subjacente EMOÇÃO É CONTÊINER, faz com que haja certas restrições quanto ao verbo infinitivo que coocorrerá. De tal modo, verbos de emoção serão preferenciais, uma vez que se sustentam pela metáfora EMOÇÃO É LÍQUIDO NUM CONTÊINER, apesar de serem possíveis ainda verbos de movimento e elocução.

Assim sendo, era necessário haver uma relação entre frames e construções que modelasse a restrição de coocorrência interna dos elementos que preenchem os ECs da *CI*. Portanto, foi definido formalmente que:

SE $V1_{fin}$ = UL que evoca, originalmente, o frame X, LOGO, $V2_{inf}$ deve evocar os frames Y, Z...

A modelagem dessa construção será realizada, sobretudo, com o auxílio do Editor de Restrições, por meio de recursos inseridos em razão da análise da *CI*. No entanto, antes de passarmos à modelagem das restrições, observe-se a definição dessa construção formalizada no Constructicon da FrameNet Brasil – Figura 46.

Auxiliar_Aspectual_Inceptiva [Inceptive_aspect_auxiliary]

Definição	
Construção aspectual em que um Auxiliar indica que o evento codificado pelo Verbo_infinitivo está em sua fase inicial. Pode haver ou não uma Preposição entre o Auxiliar e o Verbo_infinitivo .	
Exemplo(s)	
Elementos da Construção	
Auxiliar [Auxiliary]	Verbo auxiliar que indica aspecto inceptivo.
Preposição [Preposition]	Preposição que pode se interpor entre o Auxiliar e o Verbo_infinitivo .
Verbo_infinitivo [Infinitival_verb]	Verbo que codifica o evento principal.

Figura 46: Construção *Aspectual_Inceptiva*

Na Figura 46, além da definição, observam-se os ECs da *CI*, quais sejam, o *AUXILIAR*, correspondente ao $V1_{fin}$, a *PREPOSIÇÃO*, que é facultativa, e o *VERBO_INFINITIVO*, equivalente ao $V2_{inf}$. Tais informações, porém, não codificam a restrição explicitada na descrição supracitada da *CI*, de modo a não dar conta de explicar a construção por completo. Por esse motivo, o Editor de Restrições se mostra de grande relevância e aplicabilidade para o Constructicon.

Assim, a primeira restrição a ser implementada diz respeito à ordem em que os ECs devem aparecer, a chamada *CE_before*. Na *CI*, o EC *AUXILIAR* ($V1_{fin}$) deve vir antes do EC *VERBO_INFINITIVO* ($V2_{inf}$). A Figura 47 exhibe a tela gerada para se inserir tal informação.

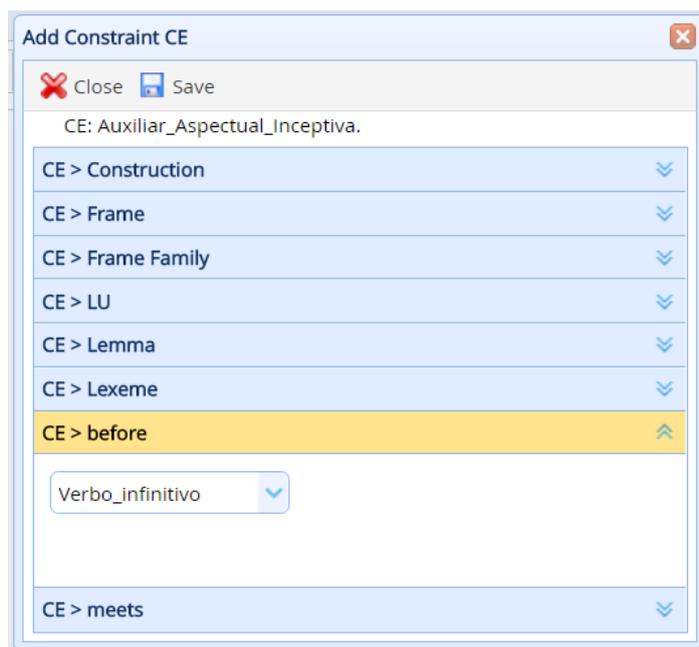


Figura 47: Restrição *CE_before*

A seguir, a restrição *CE_meets* codifica a adjacência necessária entre dois ECs. Nos casos da *CI* em que houver o EC *PREPOSIÇÃO*, nenhum material interveniente pode aparecer entre esse e o EC *VERBO_INFINITIVO* ($V2_{inf}$), conforme mostra a Figura 48.

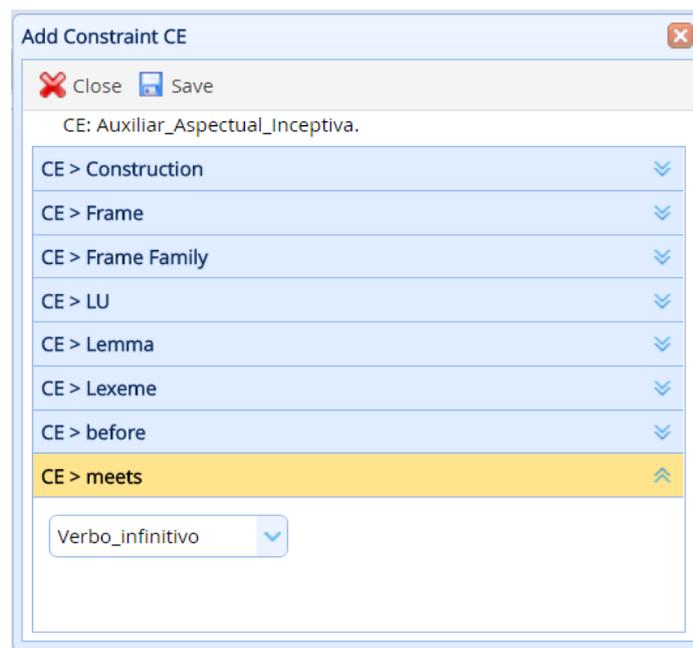


Figura 48: Restrição *CE_meets*

Outra restrição empregada na *CI* é a *CE_constructicon*, Figura 49, imputada quando uma construção já descrita no *Constructicon* licencia um dado EC. Como tanto o *AUXILIAR* ($V1_{fin}$) e o *VERBO_INFINITIVO* ($V2_{inf}$) são da classe de palavras verbo – a *FrameNet Brasil* não diferencia auxiliares e verbos plenos em termos de classes de palavras –, ambos os ECs são licenciados pela construção *Verbo*.

Vale lembrar que as três restrições apontadas, *CE_before*, *CE_meets* e *CE_construction*, foram criadas no âmbito de outro trabalho (ALMEIDA, 2016), conforme explicitado no capítulo 4. Assim, as próximas restrições a serem apresentadas são aquelas que foram instituídas por esta tese mediante a análise da *CI*.

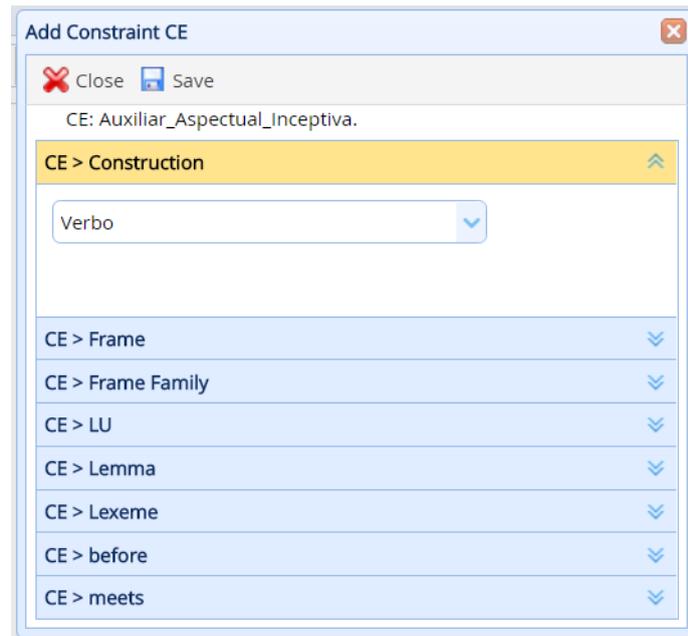


Figura 49: Restrição CE_construction

Conforme visto na análise da restrição de coocorrência, há limitações quanto ao material que pode figurar nos ECs *AUXILIAR* ($V1_{fin}$) e *VERBO_INFINITIVO* ($V2_{inf}$). Logo, foi preciso modelar novas restrições acerca das ULs que podem preencher tais ECs. Essa propriedade foi implementada em três diferentes níveis.

Em um primeiro nível, é possível restringir se apenas palavras específicas podem preencher um dado slot. Isso ocorre com o EC *AUXILIAR* ($V1_{fin}$), que só pode ser preenchido por verbos não tipicamente aspectuais específicos, a saber, “danar”, “romper”, “desatar” e “entrar”. A Figura 50 exibe essa nova restrição.

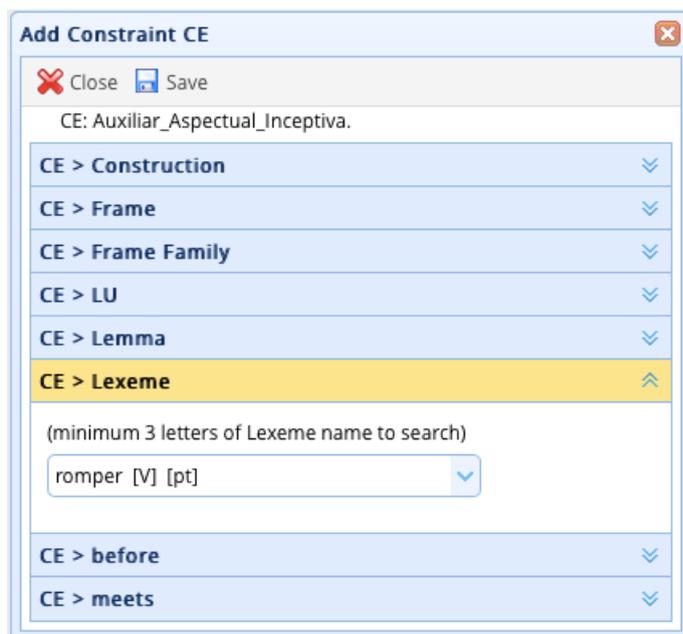


Figura 50: Restrição CE_lexeme

Em um segundo nível, há possibilidade de restringir caso todas as ULs que evocam um frame possam ocupar o dado slot, Figura 51. Se o EC *AUXILIAR* ($V1_{fin}$) for preenchido pelo verbo “romper”, todas as ULs verbais que evocam os frames de *Emoções* e *Movimento* – correspondentes aos tipos definidos por Sigiliano (2012), *emoção* e *deslocamento*, respectivamente – podem ocupar a posição do EC *VERBO_INFINITIVO* ($V2_{inf}$).

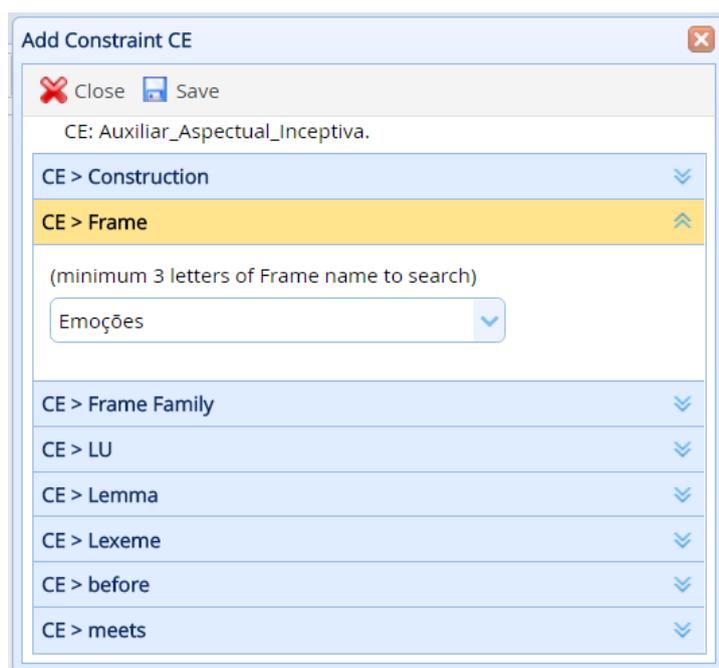


Figura 51: Restrição CE_Frame

Por fim, em um terceiro nível, inserimos a possibilidade de se restringir caso todas as ULs que evocam uma família de frames possam ocupar um determinado slot. É o que ocorre com os casos da *CI* com ‘romper’ em que o *VERBO_INFINITIVO* foi classificado por Sigiliano (2012) como sendo de ação. A FrameNet Brasil tem diversos frames que modelam ações específicas, todos eles herdeiros do frame de *Ação_transitiva*. Assim, a restrição *CE_frame_family*, Figura 52, modela que o slot do *VERBO_INFINITIVO* pode ser preenchido por ULs de todos os frames da família agentiva da FrameNet Brasil.



Figura 52: Restrição *CE_frame_family*

Aplicadas todas as restrições, a construção *Aspectual_Inceptiva* com ‘romper’ fica modelada como demonstrado pela Figura 53.

Há, ainda, um mecanismo para sinalizar caso o EC seja opcional. Esse recurso pode ser utilizado quando da criação do EC – e não no Editor de Restrições –, como explicado no capítulo 2. Na *CI*, a *PREPOSIÇÃO* é um EC cuja marcação é positiva para a propriedade opcional, já que ele pode ou não constar na sentença.

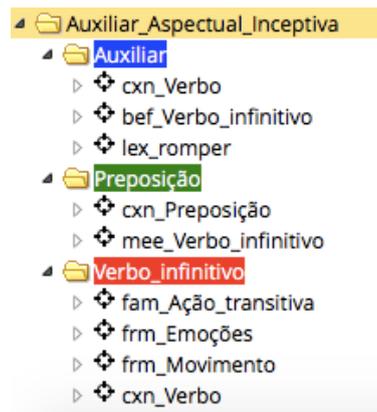


Figura 53: Restrições aplicadas à *CI*

Desse modo, a construção de *CI* é modelada no Constructicon da FrameNet Brasil, de forma que suas propriedades são todas contempladas. Ressalta-se que todas as restrições assinaladas e a marcação “opcional” descrevem a constituição da construção, ao passo que a relação de Evocação e as restrições CE_Lexeme, CE_Frame e CE_FrameFamily são responsáveis por conectar as bases de dados da FrameNet e do Constructicon relacionando frames e construções.

6 CONCLUSÕES

Na Introdução, apontamos os fatores motivadores da investigação levantada nesta tese. Primeiramente, era fundamental transformar o Constructicon em uma rede tal qual o Lexicon da FrameNet. Desse modo, buscamos na bibliografia da Gramática das Construções embasamento para estruturar a rede construcional. Aliado a isso, foi preciso considerar o viés computacional que este trabalho apresenta, fator fundamental na escolha dos parâmetros a adotar. Assim, a relação de herança, segundo Kay & Fillmore (1999), foi a escolhida, seguindo o modo completo, múltiplo e de cópia virtual.

Posteriormente, a questão de pesquisa emergiu, a qual repetimos, a seguir:

Sendo (a) a FrameNet Brasil um modelo semântico já estabelecido e (b) os itens lexicais um subtipo de construção, seria tal recurso adequado para modelar a contraparte semântica das construções descritas no Constructicon?

Ao longo desta pesquisa, buscamos responder a tal pergunta. Novamente, a literatura foi consultada a fim de que fossem verificadas as propostas de tratamento da relação entre frames e construções. Observou-se que tais construtos cognitivos, embora irmãos em sua origem, não vinham sendo sistematizados em conjunto, de modo que a maioria dos trabalhos em Gramática das Construções, ainda que considerassem os frames como responsáveis pela importação semântica a uma construção, não os tratava de maneira formalizada e menos ainda sistematizada em termos computacionais.

Buscou-se na ECG a inspiração para a criação da relação que estabeleceria o vínculo entre a construção e o frame evocado, porém, usando a base já sistematizada e referendada por anotações em *corpora* da FrameNet Brasil. Assim sendo, por meio da descrição das construções de *Dativo_com_Infinitivo* (TORRENT, 2009; LAVIOLA 2015) e da *Aspectual_Inceptiva* (SIGILIANO, 2013), foi constatada a compatibilidade entre os dois recursos, que

compõem a FrameNet Brasil – Lexicon e Constructicon –, bem como a possibilidade de modelar o aporte semântico das construções em termos de frames, sendo, portanto, positiva a resposta à pergunta inicial desta tese.

Com base, ainda, na aplicabilidade do Constructicon em tarefas de tecnologia da linguagem, buscou-se aperfeiçoar a descrição construcional, adicionando-se ao Editor de Restrições a possibilidade de delimitar a presença de uma UL, um frame, ou família de frames preenchendo um EC.

Ademais, com a implementação da relação de Evocação, o Constructicon da FrameNet Brasil destaca-se ao contribuir para o estabelecimento da interdependência entre frames e construções. Desde a origem das teorias cognitivas para a linguagem, esses dois elementos são, finalmente, formalizados, atuando em uma relação de complementaridade.

Vale destacar que nenhum Constructicon criado para outro idioma se estrutura em uma rede de heranças tal qual o recurso da FrameNet Brasil. Além disso, sinaliza-se o frame como a contraparte semântica da construção, em alguns casos – como no Constructicon do sueco –, embora não haja, efetivamente, conexão entre as bases de dados, tampouco entre os ECs e EFs.

Destarte, o Constructicon da FrameNet Brasil, a partir desta tese, avança frente a Constructica desenvolvidos para diversos idiomas, bem como em relação a outras implementações descritivas e computacionais da língua, uma vez que:

- i. converte seu recurso em uma rede construcional sustentada pela relação de herança – o que lhe confere um status superior ao de um repositório de construções;
- ii. interliga de forma sistemática as bases de dados lexical e construcional, inserindo restrições que ligam construções a frames – o que proporciona descrições mais abrangentes e completas das construções;
- iii. estabelece a união das teorias frames e construções – o que contribui para as teorias cognitivas ao revelar que tais construtos cognitivos são compatíveis e complementares, mesmo em recursos computacionais altamente sistematizados.

REFERÊNCIAS BIBLIOGRÁFICAS

ALLEN, J. **Natural Language Understanding**. New York: Pearson, 1995

ALMEIDA, V. G. **Identificação Automática de Construções de Estrutura Argumental**: um experimento a partir da modelagem linguístico-computacional das construções Transitiva Direta Ativa, Ergativa e de Argumento Cindido. Dissertação de Mestrado em Linguística. Universidade Federal de Juiz de Fora. Juiz de Fora, 2016.

BERGEN, B. **Louder than words**: The new science of how the mind makes meaning. New York: Basic Books, 2012.

_____. & CHANG, N. Embodied Construction Grammar. In: HOFFMAN, T. & TROUSDALE, G. **The Oxford Handbook of Construction Grammar** (pp.169-190). New York: Oxford University Press, 2013.

BOAS, H. C. Bilingual FrameNet Dictionaries for Machine Translation. In: RODRIGUEZ, M. G. & ARAUJO, C. P. S. **Proceedings of The Third International Conference on Language Resources and Evaluation**. Las Palmas, Espanha, 2002.

_____. A Lexical-Constructional Account of the Locative Alternation. In: CARMICHAEL, L., HUANG, & SAMIIAN, V. (orgs.). **Proceedings of the 2001 Western Conference in Linguistics**, Vol. 13, p. 27-42, 2003.

_____. Computational Resources: FrameNet and Constructicon. In: DANCYGIER, B. (org.). **The Cambridge Handbook of Cognitive Linguistics**. (pp. 549-573). Cambridge: Cambridge University Press, 2017.

BORIN, L.; DANNÉLLS, D.; FORSBERG, M.; TOPOROWSKA GRONOSTAJ, M. & KOKKINAKIS, D. Swedish FrameNet++. In: **Proceedings of the Swedish Language Technology Conference**. Linköping, 2010.

BOTELHO, L. S. Uma abordagem sociocognitiva das construções agentivas x-eiro. In: MIRANDA, N. S. & SALOMÃO, M. M. M. **Construções do Português do Brasil: da gramática ao discurso**. Belo Horizonte: UFMG, 2009.

BRYANT, J. E. **Best-Fit Constructional Analysis**. Tese de Doutorado em Ciência da Computação, University of California, Berkeley. Berkeley, b 2008.

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

_____. & CRUSE, D. A. **Cognitive Linguistics**. Cambridge: Cambridge University Press, 2004.

DIAS-DA-SILVA, B. C., MONTILHA, G., RINO, L. H. M., SPECIA, L., NUNES, M. D. G. V., DE OLIVEIRA JR, O. N. & PARDO, T. A. S. Introdução ao Processamento das Línguas Naturais e Algumas Aplicações. **Série de Relatórios do Núcleo Interinstitucional de Lingüística Computacional**, 3, 2007.

DODGE, E. **Constructional and Conceptual Composition**. Tese de Doutorado em Linguística. University of California, Berkeley. Berkeley, 2010.

FELDMAN, J. **From molecule to metaphor: a neural theory of language**. Cambridge, Mass.: A Bradford Book, 2006.

FILLMORE, C. J. The Case for Case. **Universal Linguistic Theory**, p. 1-89, 1968.

_____. The case for case reopened. **Syntax and semantics**, 8, p. 59-82, 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". In: **Annual Meeting of the Berkeley Linguistics Society**. 1988. p. 35-55.

_____. Inversion and Constructional Inheritance. In: WEBELHUTH, G., KOENIG, J. P. & KATHOL, A. **Lexical and Constructional Aspects of Linguistic Explanation**. (pp.113-128). Stanford: CSLI Publications, 1999.

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

_____. Berkeley Construction Grammar. In: HOFFMAN, T. & TROUSDALE, G. **The Oxford Handbook of Construction Grammar**. (pp. 111-132). New York: Oxford University Press, 2013.

_____, KAY, P. & O'CONNOR, M. C. Regularity and idiomaticity in grammatical constructions: The case of let alone. **Language**, p. 501-538, 1988.

_____, LEE-GOLDMAN, R. & RHOMIEUX, R. The FrameNet Constructicon. In BOAS, H. C. & SAG, I. A. (Eds.). **Sign-Based Construction Grammar**. Stanford: CSLI Publications, 2012.

_____, PETRUCK, M. R. L., RUPPENHOFER, J. & WRIGHT, A. FrameNet in Action: The Case of Attaching. **International Journal of Lexicography**, 16(3), 2003.

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

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

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

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

_____. Verbs, constructions, and semantic frames. In: RAPPAPORT-HOVAV, M., DORON, E. & SICHEL, I. (orgs.). **Syntax, lexical semantics, and event structure** (pp. 39-58). New York: Oxford University Press, 2010.

GOLDBERG, Y. & LEVY, O. word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. **arXiv preprint arXiv:1402.3722**, 2014.

HAIMAN, J. **Natural Syntax: Iconicity and Erosion**. Cambridge: Cambridge University Press, 1985.

HASEGAWA, Y., LEE-GOLDMAN, R., OHARA, K. H., FUJII, S., & FILLMORE, C. J. On expressing measurement and comparison in English and Japanese. In: BOAS, H. C. **Contrastive construction grammar** (pp. 169-200). Amsterdam: John Benjamins, 2010.

HEINE, B. & KUTEVA, T. **Language contact and grammatical change**. Cambridge: Cambridge Univ. Press, 2005.

HEINE, B. & KUTEVA, T. **The changing languages of Europe**. Oxford: Oxford University Press, 2006.

IBÁÑEZ, F. J. R. M. & MASEGOSA, A. G. **Cognitive Modeling: A linguistic perspective**. Amsterdam: John Benjamins, 2014.

KATZ, J. J. **Semantic Theory**. New York: Harper and Row, 1972.

KAY, P. Argument structure constructions and the argument-adjunct distinction. In: FRIED, M. & BOAS, H. C. **Grammatical constructions: Back to the roots.** (pp. 71-98). Amsterdam: John Benjamins, 2005.

_____. & FILLMORE, C. J. Grammatical Constructions and Linguistic Generalizations: the What's X Doing Y Construction. **Language**, vol. 75, nº 1, 1999.

KOEHN, P. **Statistical machine translation.** Cambridge: Cambridge University Press, 2009.

LAGE, L. M. **Frames e construções: A implementação do constructicon na FrameNet Brasil.** Dissertação (Mestrado em Linguística). Universidade Federal de Juiz de Fora, Juiz de Fora, 2013.

LAKOFF, G. & BRUGMAN, C. The semantics of aux-inversion and anaphora constraints. In: **Linguistic Society of America Meeting**, San Francisco, CA. 1987.

LAKOFF, G. & JOHNSON, M. **Metáforas da Vida Cotidiana.** Campinas, SP: Mercado de Letras, 2002.

LAMBRECHT, K. **Information structure and sentence form: Topic, focus, and the mental representations of discourse referents.** Cambridge: Cambridge University Press, 1994.

LAVIOLA, A. B. & 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.

LAVIOLA, A. B. **Frames e Construções em Contraste: uma análise comparativa Português - Inglês no tangente à implementação de Constructicons.**

Dissertação de Mestrado em Linguística. Universidade Federal de Juiz de Fora. Juiz de Fora, 2015.

LEE-GOLDMAN, R. & PETRUCK, M. R. L. **The FrameNet Constructicon in Action**. (no prelo).

LYNGFELT, B.; BORIN, L.; FORSBERG, M.; PRENTICE, J.; RYDSTEDT, R.; SKÖLDBERG, E. & TINGSSELL, S. Adding a Constructicon to the Swedish resource network of Språkbanken. In: **Proceedings of KONVENS 2012 (LexSem 2012 workshop)**, pp. 452-461. Vienna, 2012.

LYNGFELT, B.; BÄCKSTRÖM, L.; BORIN, L.; EHRLEMARK, A. & RYDSTEDT, R. Constructicography at work: Theory meets practice in the Swedish Constructicon. In: BORIN, L.; LYNGFELT, B.; OHARA, K. & TORRENT, T. T. **Constructicography: Constructicon Development across Languages**. Amsterdam: John Benjamins, no prelo.

MATOS, E. E.; TORRENT, T. T.; ALMEIDA, V. G.; SILVA, A. B. L.; LAGE, L. M.; MARCAO, N. D. & TAVARES, T. S. Constructional Analysis Using Constrained Spreading Activation in a FrameNet-Based Structured Connectionist Model. In: **The AAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02**. Palo Alto, CA: AAI Publications, v.17, p.222-229, 2017.

MARÇÃO, N. D.; TORRENT, T. T. & MATOS, E. E. S. Descrição e modelagem de construções interrogativas QU-em Português Brasileiro para o desenvolvimento de um chatbot. In: **Proceedings of the 11th Brazilian Symposium in Information and Human Language Technology**, p. 209-216, 2017.

MICHAELIS, L. A. & LAMBRECHT, K. The exclamative sentence type in English. In: GOLDBERG, A. E. (org.) **Conceptual Structure, discourse and language**, p. 375-389, 1996.

MICHAELIS, L. A. & RUPPENHOFER, J. **Beyond Alternations: A Constructional Model of the German Applicative**. Stanford: CSLI Publications, 2001.

MINSKY, M. A framework for representing knowledge. In: WINSTON, P. (org.). **The psychology of computer vision**. New York: Mc Graw-Hill, 1975.

NARAYANAN, S. **Knowledge-Based Action Representations for Metaphor and Aspect (KARMA)**. Tese de Doutorado em Ciência da Computação. University of California, Berkeley. Berkeley, 1997.

OHARA, K. H. Toward Constructicon Building for Japanese in Japanese FrameNet. **Veredas (UFJF. Online)**, v. 17, p. 11-27, 2013.

_____.; FUJII, S.; OHORI, T.; SUZUKI, R.; SAITO, H. & ISHIZAKI, S. The Japanese FrameNet Project: An Introduction. In: **Proceedings of the LREC 2004 Satellite Workshop: Building Lexical Resources from Semantically Annotated Corpora**. Lisboa, 2004.

PARTEE, B. H. Subject and Object in Modern English. In: HANKAMER, J. (org.) **Outstanding Dissertations in Linguistics Series**. New York: Garland, 1979 [1965].

PETRUCK, M. R. L. & DE 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.

PINKER, S. **Learnability and Cognition: The Acquisition of Argument Structure**. Cambridge, Mass.: MIT Press, 1989.

PUSTEJOVSKY, J. & STUBBS, A. **Natural Language Annotation for Machine Learning**: A guide to corpus-building for applications. New York: O'Reilly Media, 2012.

RUPPENHOFER, J.; ELLSWORTH, M.; PETRUCK, M. R. L.; JOHNSON, C. R. & SCHEFFCZYK, J. **FrameNet II: Extended theory and practice**. Berkeley: International Computer Science Institute, 2016.

RAPPAPORT-HOVAV, M.; LEVIN, B. Reflections on Manner/Result Complementarity. In: RAPPAPORT-HOVAV, M., DORON, E. & SICHEL, I. (orgs.). **Syntax, lexical semantics, and event structure** (pp. 21-38). New York: Oxford University Press, 2010.

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.

_____. 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. 2007.

_____. Construções modais com dar no português do Brasil: metáfora, uso e gramática. **Revista de estudos da linguagem**, v. 16, n. 1, p. 83-115, 2008.

_____. 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, 2009a.

_____. FrameNet Brasil: um trabalho em progresso. **Calidoscópico**, v. 7, n. 2, p.171-182, 2009b.

_____., TORRENT, T. T. & SAMPAIO, T. F. A Linguística Cognitiva Encontra a Linguística Computacional: notícias do projeto FrameNet Brasil. **Cadernos de Estudos Linguísticos**, 53(2), 2013.

SIGILIANO, N. S. Persistência e extensão na construção inceptiva em [V1fin (Prep) V2inf] do português. **Letras & Letras**, Uberlândia, Universidade Federal de Uberlândia, Instituto de Letras e Linguística, vol. 27, nº 1, jan./jun. 2012.

_____. A Construção Aspectual Inceptiva do Português com Verbos Não Canônicos. Tese de Doutorado em Linguística. Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2013.

SPÄRCK-JONES, K. Computational linguistics: what about the linguistics? **Computational linguistics**, v. 33, n. 3, p. 437-441, 2007.

SUBIRATS, C. & PETRUCK, M. R. L. Surprise: Spanish FrameNet! In: **Proceedings of the 17th International Conference of Linguists**. Prague, 2003.

TALMY, L. **Toward a Cognitive Semantics**. Vol. I. Cambridge, Massachusetts: The MIT Press, 2001.

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. (orgs.). **LETRAS & LETRAS**. Uberlândia: UFU, Volume 27 - Número 1 - jan./jun.. 2011.

_____. On the relation between inheritance and change: The Constructional Convergence and the Construction Network Reconfiguration Hypotheses. In: BARDDAL, J.; SMIRNOVA, E.; SOMMERER, L. & GILDEA, S. **Diachronic Construction Grammar**. (pp. 174-211). Amsterdam: John Benjamins, 2015.

_____. & ELLSWORTH, M. J. Behind the Labels: Criteria for Defining Analytical Categories in FrameNet Brasil. **Veredas (UFJF. Online)**, v. 17, p. 44-65, 2013.

_____.; MATOS, E. E.; LAGE, L. M.; LAVIOLA, A. B.; TAVARES, T. S. & ALMEIDA, V. G. Towards the Continuity Between the Lexicon and the Constructicon in FrameNet Brasil. In: BORIN, L.; LYNGFELT, B.; OHARA, K. & TORRENT, T. T. **Constructicography: Constructicon Development across Languages**. Amsterdam: John Benjamins, no prelo.

YOU, L & LIU, K. Building Chinese FrameNet Database. In: **Proceedings of the 2005 IEEE International Conference on Natural Language Processing and Knowledge Engineering**. Shanghai: ICCS Shanghai Normal University, 2005.

ZIEM, A. & BOAS, H. C. Towards a Constructicon for German. In **The AAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02**. Palo Alto, CA: AAI Publications, v.17, p. 274-277, 2017.