

REPRESENTAÇÃO COMPUTACIONAL DAS CONSTRUÇÕES DE SUJEITO- PREDICADO DO PORTUGUÊS DO BRASIL

Alexandre Diniz da Costa (UFJF)¹

Vânia Gomes Almeida (UFJF)²

Ludmila Meireles Lage (UFJF)³

Gustavo Barbosa (UFJF)⁴

Natália Duarte Marção (UFJF)⁵

Vanessa Ramos Lopes Paiva (UFJF)⁶

Ely Edison da Silva Matos (UFJF)⁷

Tiago Timponi Torrent (UFJF)⁸

RESUMO

Este trabalho apresenta uma proposta para representar computacionalmente construções do Português Brasileiro, no âmbito do Constructicon da FrameNet Brasil. Dessa forma, demonstra de que maneira as teorias irmãs da Semântica de Frames e da Gramática das Construções podem ser

1 Doutorando do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: alexandre.costa@ufjf.edu.br

2 Doutoranda do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: vania.almeida2017@letras.ufjf.br

3 Doutoranda do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: lu.meireleslage@gmail.com

4 Mestrando do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: gustavo.grpb@gmail.com

5 Mestranda do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: duarte.natalia@letras.ufjf.br

6 Mestranda do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: vanessalettrasufjf@gmail.com

7 Professor colaborador do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: ely.matos@ufjf.edu.br

8 Professor Adjunto do Programa de Pós-graduação em Linguística da Universidade Federal de Juiz de Fora. E-mail: tiago.torrent@ufjf.edu.br

implementadas computacionalmente com vias a sustentar aplicações em Compreensão de Língua Natural.

Palavras-chave: Representação Computacional. Gramática de Construções. Semântica de Frames. FrameNet Brasil.

ABSTRACT

This paper presents a proposal for the computational representation of constructions in Brazilian Portuguese in the FrameNet Brasil Constructicon. We demonstrate how the sister theories of Frame Semantics and Construction Grammar can be implemented so as to support applications in Natural Language Understanding.

Keywords: Computational Representation. Construction Grammar. Frame Semantics. FrameNet Brasil.

Introdução⁹

Implementações computacionais de Gramáticas das Construções têm ganhado cada vez mais espaço no campo da Linguística Computacional dedicado à Compreensão de Língua Natural (Allen, 1995). Tais implementações são aplicadas às mais diversas tarefas, tais como *parsing* sintático-semântico (Bryant, 2008; Steels, 2011; Marques & Beuls, 2016; Matos et al., 2017), comunicação com sistemas autônomos (Trott et al., 2015) e geração automática de língua natural (Dominey et al., 2017).

Todas essas aplicações dependem, em alguma medida, de um constructicon, ou seja, um recurso computacional em que informações estruturais sobre construções de uma ou mais línguas são modeladas (cf. Lyngfelt, no prelo). Um recurso dessa natureza encontra-se atualmente em desenvolvimento para o Português Brasileiro (PB), no âmbito do projeto FrameNet Brasil, cujo objetivo principal é o de construir, para esse idioma, um modelo semântico baseado em frames, atrelado a suas manifestações linguísticas (Salomão, 2009).

9 O Constructicon da FrameNet Brasil é desenvolvido no âmbito do projeto de pesquisa Frames e Construções em Contraste, financiado pelo programa CAPES/STINT de cooperação binacional Brasil-Suécia (Processo nº 99999.009910/2014-00).

Dado esse contexto, o objetivo principal deste trabalho é o de apresentar uma proposta de representação computacional de construções do PB, fundada na Semântica de Frames (Fillmore, 1982; 1985) e na Gramática de Construções (Fillmore, Kay & O'Connor, 1988; Kay & Fillmore, 1999; Fillmore, 2013), implementada no Constructicon da FrameNet Brasil. Para tanto, na seção 1, apresentamos uma breve contextualização da interface entre representações computacionais e linguística cognitiva, apresentando, na sequência – seção 2 – a FrameNet Brasil como uma proposta de representação de conhecimento linguístico baseada na Semântica de Frames. Na seção 3, demonstraremos como a proposta de modelagem objeto deste trabalho é implementada, através da representação dos aportes formal e funcional de construções de estrutura argumental no PB. Na seção 4, discutimos as limitações existentes no modelo, para, na seção 5, apresentarmos nossas conclusões.

1. A Interface entre Representações Computacionais e Linguística Cognitiva

Poole e Mackworth (2010), ao tratar do desenvolvimento de agentes computacionais baseados em Inteligência Artificial (IA), definem as características fundamentais das representações computacionais de conhecimento nos seguintes termos:

uma representação deve ser suficientemente rica para expressar o conhecimento necessário para resolver o problema. Ela deve ser o mais próximo possível do problema. Deve ser compacta, natural e sustentável. Deve ser fácil ver a relação entre a representação e o domínio que está sendo representado, de modo que seja fácil determinar se o conhecimento representado está correto. Uma pequena mudança no problema deve resultar em uma pequena mudança na representação do problema. Essa representação deve ser acessível a uma computação eficiente, o que geralmente significa que é capaz de expressar características do problema que podem ser exploradas para ganho computacional e capaz de balancear precisão e tempo de computação. Ela deve poder ser obtida a partir de pessoas, dados e experiências passadas. (Poole & Mackworth, 2010)¹⁰

Tais traços podem, em princípio, ser aplicáveis a qualquer tipo de conhecimento e/ou capacidade cognitiva que se queira representar computacionalmente, seja ela visão, reconhecimento de fala ou a representação do conhecimento linguístico em seus aspectos formais e semântico-pragmáticos.

No âmbito da Linguística, a busca por representações ocasionou o surgimento de inúmeros

10 "A representation should be rich enough to express the knowledge needed to solve the problem. It should be as close to the problem as possible; it should be compact, natural, and maintainable. It should be easy to see the relationship between the representation and the domain being represented, so that it is easy to determine whether the knowledge represented is correct. A small change in the problem should result in a small change in the representation of the problem. It should be amenable to efficient computation, which usually means that it is able to express features of the problem that can be exploited for computational gain and able to trade off accuracy and computation time. It should be able to be acquired from people, data and past experiences."

formalismos de representação do conhecimento, tais como a semântica formal, a lógica formal, os sistemas de frames, as redes semânticas e a gramática de construções, entre outros.

Teorias como o Gerativismo Chomskyano (Chomsky, 1993 [1981]) se destacaram, nesse sentido, justamente por considerar a manipulação do sentido através do estudo da forma, ou seja, por tratar o significado como imanente. Nessas perspectivas, de modo genérico, a proposta era a de que o léxico forneceria um mecanismo que asseguraria que todo falante de português, por exemplo, no emprego de um verbo como *amanhecer*, não selecionasse para ele qualquer argumento interno ou sujeito [+animado].

Ao se propor realizar representações computacionais, não se pode negar que é muito mais cômodo manipular um modelo de língua ideal, que não apresenta idiossincrasias ou usos metafóricos e metonímicos. No entanto, não podemos ignorar que aspectos como esses estão no cerne do nosso processamento cognitivo como operações fundamentais, sob pena de terminarmos com um modelo computacional “inocente”, que não consegue dar conta nem mesmo dos usos mais cotidianos da língua, processando o significado das sentenças de acordo com as suas partes e a sua organização, numa extensão, para o domínio da modelagem, dos termos de Fillmore (1979). Para este autor, o falante-ouvinte inocente, apesar de conhecer as regras sintáticas da língua, não tem conhecimento sobre expressões idiomáticas ou usos metafóricos, processando a língua literal e composicionalmente.

Sentenças como (1) não satisfariam as condições necessárias, segundo um modelo baseado em regras, para a adequada interpretação do verbo em questão, uma vez que o verbo *amanhecer*, nesse caso, tem sujeito humano e complemento predicativo. Entretanto, qualquer falante do português é capaz de compreender tal sentença.

(1) O homem amanheceu doente.

Face a ocorrências como (1), uma abordagem lexicalista, que confiasse exclusivamente nas propriedades de seleção argumental do verbo para a organização da sentença, seria levada a propor a existência de uma outra entrada no léxico mental para *amanhecer*, a qual seria homônima daquela descrita anteriormente, mas que teria, em sua grade temática, um sujeito [+animado] e seu predicativo.

A Linguística Cognitiva surge com uma proposta de tratamento diferenciado para o significado, uma vez que a relação entre as formas linguísticas e o sentido atribuído a elas é mediada pela cognição.

O significado deixa de ser um reflexo direto do mundo e passa a ser visto como uma construção através da qual o mundo é apreendido e experienciado. Assim, a forma não contém o significado, mas contribui para a sua construção (Fauconnier & Turner, 2002).

Nesse contexto, a Semântica de Frames se destaca como um programa de pesquisa em semântica que enfatiza a relação entre língua e experiência, conceptualizando o mundo em cenas ou frames. Nas palavras de seu fundador, o linguista Charles Fillmore, um frame é “um sistema de conceitos relacionados de tal forma que, para entender qualquer um desses conceitos, é necessário entender todo o sistema em que estão inseridos”¹¹ (Fillmore, 1982, p.111). Assim, para que se entenda a distinção de sentido entre (2) e (3), bem como a estranheza que um professor causaria a seus alunos universitários ao dizer (2), não basta conhecer o significado “de dicionário” das palavras café e merenda. Faz-se necessário a ativação de uma estrutura de conceitos – ou frame – que caracteriza um estabelecimento de ensino como tendo um determinado perfil de alunos, os quais pertencem a uma dada faixa etária.

(2) Meninos, já está na hora da merenda, podem sair.

(3) Meninos, já está na hora do café, podem sair.

Tal definição de frame e suas implicações para o estudo do significado linguístico relacionam-se, desde sua proposição, ao conceito de frames em IA (Minsky, 1975), donde deriva a natural aproximação entre a teoria da Semântica de Frames e as representações computacionais. Minsky (1975) propõe que os segmentos de raciocínio, linguagem, memória e percepção sejam maiores e mais estruturados do que meras listas de itens e suas definições. Em sua proposta teórica para a representação de conhecimento no âmbito da IA, parte do pressuposto de que

Quando alguém se depara com uma situação nova (ou causa uma mudança substancial em sua percepção do problema em questão), essa pessoa seleciona na memória uma estrutura chamada Frame. Trata-se de um modelo armazenado na memória que deve ser adaptado à realidade através da mudança de detalhes, caso seja necessário. Um frame é uma estrutura de dados para representar uma situação estereotipada, como estar em um tipo de sala de visitas, ou ir a um aniversário de criança.¹²

11 "a system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits"

12 When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary. A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party.

Minsky (1975) endossa a visão, que começa a tomar forma em Fillmore (1968), de que a significação linguística não se constitui da soma dos significados dos itens em uma sentença, mas que confia em algum tipo de estrutura mais ampla que organiza o pano de fundo para a compreensão do sentido. Na altura de publicação de sua proposta teórica, Minsky se refere à Gramática de Casos, teoria fillmoreana que se constitui na semente da Semântica de Frames e de sua teoria irmã, a Gramática de Construções.

A vertente da Gramática de Construções desenvolvida por Fillmore e Kay em Berkeley (Kay & Fillmore, 1999; Fillmore, 2008; 2013), constitui, portanto, outra das aproximações entre a Linguística Cognitiva e as representações computacionais. Caracterizada por ser um modelo baseado em unificação que se fundamenta matematicamente, busca incorporar práticas computacionais para tratar o significado das expressões linguísticas. Nesse contexto, Fillmore, Kay e O'Connor (1988) apresentam construções como as regras que licenciam signos linguísticos baseados em outros signos linguísticos. As estruturas licenciadas por uma ou mais construções são chamadas de construtos, os quais instanciam propriedades particulares com respeito a uma construção que os licencia.

Através da unificação entre os níveis de representação, a Berkeley Construction Grammar apresenta uma relação estreita entre significado e significante por um sistema baseado em traços em que as construções são representadas por Matrizes de Atributo e Valor (AVM), em um pareamento de forma e sentido mediado por signos.¹³ As AVMs são combinadas de acordo com os valores atribuídos aos seus atributos e a coindexação mantém o controle da unificação de relações.

Assim, uma sentença como (1), condenada a ser ou mal formada, ou excepcional, ou periférica em outras abordagens sobre a gramática, pode ser formalizada conforme a AVM na Figura 1, a qual representa o construto licenciado pela Construção de Mudança de Estado do PB. Tal construção é um dos subtipos de construção de Sujeito-Predicado do PB, sendo composta por dois signos filhos: um Sintagma Nominal Determinado, em posição de sujeito, e um Sintagma Verbal Adjetival, na função de predicado, configurando-se o esquema sintático [[SN[V [SA_{dj}]]]. Do ponto de vista semântico, codifica o frame *Sofrer_mudança*, no qual uma Entidade, o Sujeito, passa por uma mudança de estado, até atingir uma *Qualidade_final*, codificada pelo Adjetivo em função de Predicativo, o qual

¹³ Vide a Entrevista com Paul Kay, neste volume, para uma discussão sobre a relação entre a Berkeley Construction Grammar e a Sign-Based Construction Grammar.

evoca o frame de Condições_médicas. As Circunstâncias em que a mudança se dá são marcadas pelo Verbo, que evoca o frame de Mudança_de_turno, no qual uma Unidade calêndrica avança no tempo, passando de um turno a outro.

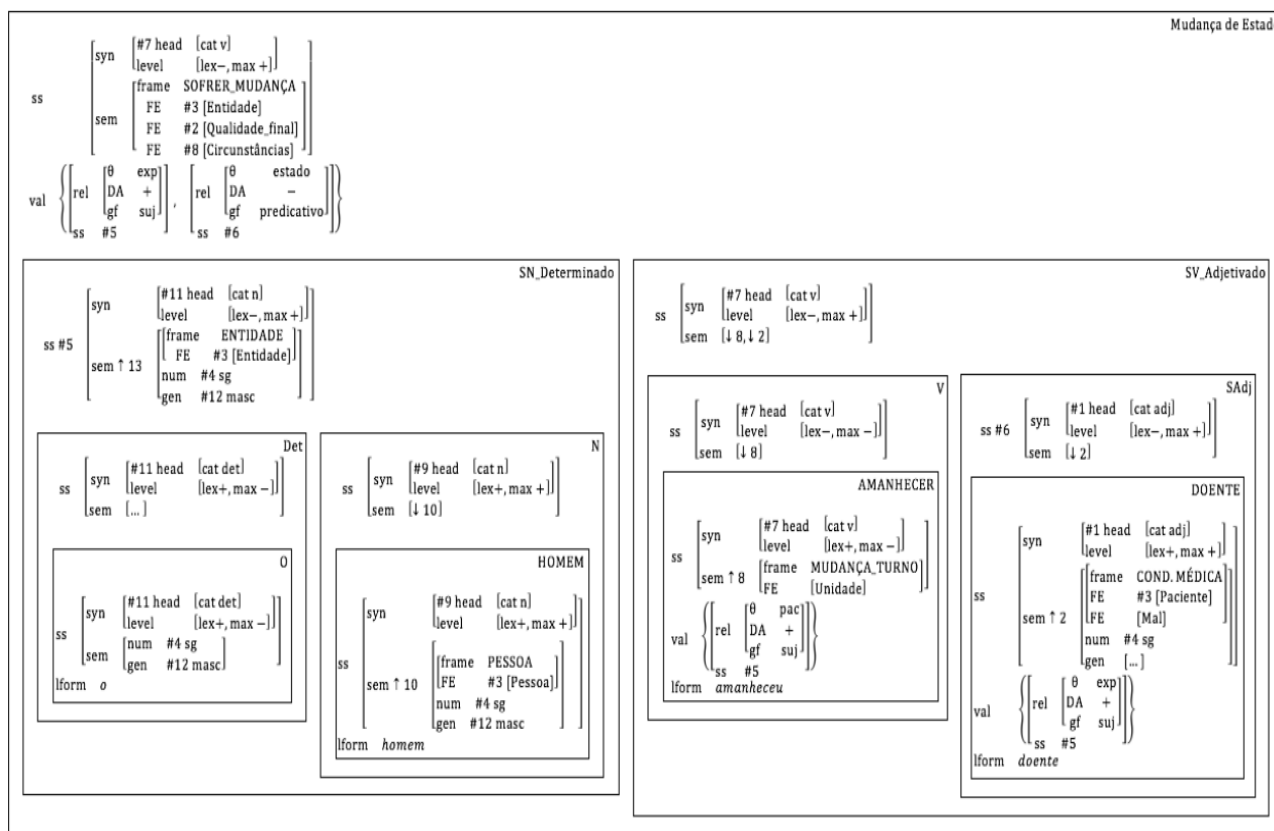


Figura 1: AVM do Construto “O homem amanheceu doente”, licenciado pela Construção de Mudança de Estado

Na AVM, cada constituinte é representado por uma caixa, que contém as seguintes estruturas básicas:

- Nome do constituinte: aparece no canto superior direito;
- Estrutura sintático-semântica (ss): conjunto de atributos e seus valores que indicam as propriedades sintáticas (syn) e semânticas (sem) de cada constituinte e/ou da construção como um todo. É composta pelos seguintes atributos:
 - head: indica o traço de núcleo do constituinte, o qual se define quanto ao pertencimento categorial (cat), ou seja, se se trata de um nome (n), verbo (v), adjetivo (a), determinante (det) etc., e ao nível (level), isto é, se se trata de um constituinte que é uma projeção máxima de seu núcleo (max +/-) e se este tem manifestação lexical direta (lex +/-);
 - frame: indica o frame evocado pelo constituinte, o qual é definido por seu nome e

pelos elementos de frame (FEs) que o compõem;

- num: indica se o constituinte é singular ou plural;
 - gen: indica se o constituinte é masculino ou feminino.
- Valência (val): representa a valência do constituinte e/ou da construção como um todo. Cada valente é definido quanto à sua estrutura ss e também quanto à relação (rel) que estabelece com o constituinte, tanto em termos das funções teta (θ) e gramatical (gf) a ele atribuídas, quanto ao fato de ser ou não um argumento semanticamente proeminente (DA +/-).

Os índices numéricos presentes na matriz indicam a unificação dos valores de cada atributo. O processo de unificação pode indicar dois aspectos importantes da construção. Por um lado, checa se os atributos sintático-semânticos dos constituintes estão em acordo. Assim, os índices #4 e #12, por exemplo, dão conta da necessária concordância nominal entre o sujeito e seu predicativo, bem como daquela entre o determinante e o nome. Por outro, parecia forma e sentido, ao mapear a semântica de cada elemento presente na construção. Assim, o índice #3 indica que a Entidade que sofre a mudança de estado é uma Pessoa, que pode ser enquadrada como o Paciente do frame de Condições_médicas. Já o índice #2 indica que o estado final de #3 é estar doente, enquanto #8 mostra que a Circunstância em que tal mudança se deu corresponde à passagem da noite para a manhã. Índices precedidos de setas indicam que determinado aspecto semântico é projetado para um constituinte maior.

Por fim, o atributo da forma lexical (lform) indica quais elementos da sentença que contêm o construto licenciado estão mapeados a quais constituintes.¹⁴

A representação de construções na forma de AVMs, típica da BCG e da Sign-Based Construction Grammar – SBCG – (Sag, 2012), aliada a uma representação semântica baseada em frames, abre caminho para a modelagem computacional de construções, em especial dada a existência de recursos já implementados com larga trajetória de desenvolvimento, como a FrameNet. Na próxima seção, apresentamos esse recurso.

2. A FrameNet e o Constructicon

Em 1997, Fillmore e colaboradores passam a desenvolver uma implementação computacional de sua teoria: a Berkeley FrameNet tem início no International Computer Science Institute, com o objetivo de descrever lexicograficamente a língua inglesa a partir de uma modelagem semântica

¹⁴ Para uma explicação detalhada dos atributos comumente constantes de AVMs, vide Fried & Ostman (2004).

fundada em uma rede de frames (Fillmore et al., 2003). Nesse recurso, Unidades Lexicais (LUs) são descritas a partir de um frame de *background*, e instâncias dessas LUs são anotadas em sentenças extraídas de *corpora*. Os frames, por sua vez, são organizados em uma rede e definidos em termos de seus participantes, denominados Elementos de Frame (FEs). A Figura 2 traz, como exemplo, o frame *Medical_conditions*.

Note-se que o frame traz uma definição pensada para os usuários humanos do recurso, a qual vem seguida dos FEs que compõem o frame, com suas respectivas definições. Os FEs são classificados em nucleares (core) e não-nucleares (non-core). Enquanto aqueles são necessários para que o frame se instancie, estes são opcionais.

[Lexical Unit Index](#)

Medical_conditions

Definition:

Words in this frame name medical conditions or diseases that a patient suffers from, is being treated for, may be cured of, or die of. The condition or disease may be described in a variety of ways, including the part or area of the body (**Body part**) affected by the condition (e.g. liver cancer, cardiovascular disease), the **Cause** of the condition (e.g. bacterial meningitis, viral pneumonia), a prominent **Symptom** of the condition (e.g. asymptomatic stenosis, blue ear disease), the **Patient** or population (originally) affected by the condition (e.g. bovine tuberculosis, juvenile diabetes), or the (proper) **Name** used to identify the condition (e.g. Munchausen Syndrome, Lou Gehrig's Disease). Annotation in this frame is done in respect to the name of the condition or disease.

John died of **pancreatic** **CANCER**.

FEs:

Core:

Ailment [I] FN: Any medical problem.

Patient [Pat] This FE is used for the patient, entity, or population (originally) affected by the condition or illness.
Semantic Type: Living_thing
juvenile **DIABETES**
Dutch Elm **DISEASE**
Legionnaire's **DISEASE**

Non-Core:

Body part [Body_P] Body Part is used for the part or area of the body affected by the condition or disease.
cardiovascular **DISEASE**
liver **DISEASE**

Cause [Cause] This FE is used for the cause of a condition or disease.
bacterial **MENINGITIS**

Degree [deg] The extent to which the **Ailment** diverges from the norm.
Semantic Type: Degree
I am **very** **SICK**.
Since the degree of involvement with an **Ailment** is equivalent to how negative the experience of the **Patient** is, in many cases, **Degree** is expressed by a negative descriptor.
It seems that he does have **very severe** **ECZEMA**.

Duration [dur] The length of time over which a medical condition persists.

Name [Name] This FE is used for the **Name** that identifies the condition or disease.
Munchausen **SYNDROME**

Place [I] The location where the **Ailment** is effecting the **Patients**.

Symptom [Symp] This FE is used for a prominent symptom of the condition or disease.
asymptomatic **STENOSIS**
I have a **COLD** with a really bad cough.

Figura 2: Frame de *Medical_conditions*, da Berkeley FrameNet

Uma vez definido o frame, LUs que o evocam podem ser anotadas tanto para os FEs que aparecem em sua localidade sintática, quanto para as funções gramaticais (GFs) e tipos sintagmáticos (TS) desses FEs.

A partir do projeto matriz, outros grupos de pesquisadores passaram a desenvolver framenets para seus idiomas, dentre os quais o PB, através da FrameNet Brasil (Salomão, 2009; Torrent & Ellsworth, 2013). Assim, a sentença em (1) pode ser anotada para a LU *doente.a* do frame Medical_conditions (ou Condições_médicas, em PB), conforme a Figura 3.

	NI	0	homem	amanheceu	doente.
FE	INI		Pacient		
GF			Ext		
PT			NP		
Other					
Adj				Cop	

Figura 3: Anotação lexicográfica na FrameNet Brasil

Note-se que, na anotação lexicográfica, o SN “o homem”, que instancia o FE Paciente, é marcado com a GF de Argumento Externo (Ext) do adjetivo, ao qual se liga por uma cópula, indicada na camada Adj. O FE Mal, necessário à compreensão do frame, é marcado como uma Instanciação Nula Indefinida (INI), uma vez que não se manifesta linguisticamente, mas é inferível, uma vez que, se o homem está doente, é porque foi acometido de algum Mal (cf. Fillmore, 2007, para uma discussão sobre instanciações nulas na FN).

A anotação mostrada na Figura 3, entretanto, não é capaz de capturar a semântica da construção de Mudança de Estado que licencia a sentença (1). Isso porque ela se dá a partir da LU *doente.a*. Devido à constatação dos limites impostos pela anotação lexicográfica, a Berkeley FrameNet passa a desenvolver, em 2008, em paralelo ao já estruturado Lexicon, um Constructicon, com o intuito de representar computacionalmente determinadas estruturas linguísticas não processáveis lexicograficamente (Fillmore, 2008; Fillmore, Lee-Goldman & Rhomieux, 2012). O Constructicon abarca o conhecimento linguístico que excede a valência simples de palavras simples (Fillmore, 2008). De modo mais específico, descreve construções em termos de suas propriedades gramaticais e seu potencial semântico.

A iniciativa de criação de um modelo computacional de construções foi seguida também pela

FrameNet Brasil que, desde 2010, desenvolve um Constructicon para o PB (Torrent et al., 2014; Silva et al., 2017). Da mesma forma que acontece para unidades lexicais, o Constructicon também realiza anotações, porém enquanto o método lexicográfico representa apenas a LU e os seus elementos de acordo com seu padrão de valência, o método construcional mapeia uma construção formalmente, unificando-a, quando relevante, a um frame específico.

Assim, para que a sentença (1) seja corretamente analisada, faz-se necessário a modelagem, no Constructicon da FrameNet Brasil, da Construção de Mudança de Estado. O processo de representação computacional dessa e de outras construções de Sujeito-Predicado do PB será objeto da próxima seção.

3. Modelagem das Construções de Sujeito-Predicado no Constructicon da FrameNet Brasil

As construções analisadas neste trabalho são herdeiras da construção de Sujeito-Predicado, a qual, por sua vez é um dos subtipos da construção X_Núcleo. Tal construção abstrata é composta por um núcleo que é especificado por um elemento à esquerda (Kay & Fillmore, 1999). Serão mostradas nas próximas subseções, as representações computacionais das construções Transitiva_direta_ativa, Intransitiva, Ergativa, Predicativa, de Mudança de Estado, de Argumento Cindido (Sampaio, 2010) e de Objeto Interdito (Bronzato, 2009). Antes, entretanto, apresentamos os mecanismos básicos utilizados para representar, no domínio computacional, os princípios da constituência e da unificação empregados na constituição de AVMs. Tais mecanismos foram desenvolvidos no âmbito das pesquisas de Almeida (2016) e de Lage (2018).

Nesse contexto, a fim de conferir ao Constructicon o *status* de rede construcional, além da possibilidade de armazenar informações de modo econômico e não redundante, foi estabelecida a relação de Herança construcional (Lage, 2018). É adotada a herança múltipla, modelo segundo o qual uma construção, na hierarquia, pode herdar de mais de uma construção dominante. Além disso, o modo completo de herança é utilizado (Kay & Fillmore, 1999), o qual presume que a construção herdeira é um tipo mais específico da construção mãe, uma vez que toda informação específica para cada nó que domina, direta ou indiretamente, um determinado nó é herdada (Goldberg, 1995, p. 74). Ainda, a herança será tratada pelo viés da cópia virtual de informação, isto é, a informação herdada é registrada somente na construção dominante, o que impede a redundância no armazenamento de dados.

Por outro lado, com o intuito de representar a unificação entre os polos formal e funcional das

construções, foi estabelecida a relação de Evocação (Lage, 2018), responsável por associar a construção ao frame por ela evocado, nas situações em que isso ocorrer. Ao fazê-lo, são também relacionados os Elementos da Construção (CEs) aos Elementos de Frame (FEs), quando houver tal correspondência.

Por fim, para contemplar a constituição construcional, foi necessário implementar restrições que envolvessem uma série de aspectos. Em um primeiro momento, foram criadas três restrições, a saber, a CE>Constructicon, a CE_before e a CE_meets (Almeida, 2016). A restrição CE>Constructicon foi implementada para determinar os casos em que o signo filho é licenciado por outra construção já definida no Constructicon. Por sua vez, a CE_before determina se um dado Elemento da Construção deve vir antes de outro na sentença, enquanto a CE_meets sinaliza que um Elemento da Construção deve vir antes de outro e que ambos devem ser adjacentes, de modo que não haja entre eles nenhum material interveniente.

Posteriormente, outras três restrições foram elaboradas, quais sejam, CE>Lexeme, CE>Frame e CE>Frame_family, instanciando três níveis de restrições de preenchimento. No primeiro deles, restringe-se que apenas palavras específicas possam preencher um dado *slot*. No segundo nível, é possível restringir-se que todas as LUs que evocam um frame possam ocupar um dado *slot*. No terceiro nível, pode-se restringir que todas as LUs que evocam uma família de frames, isto é, um dado frame mais genérico e todos os frames que herdaram dele, possam ocupar um dado *slot*.

Exemplos da aplicação das relações e restrições são apresentados nas próximas seções.

3.1 A Construção de Mudança de Estado

Conforme já pontuado na seção 2, a construção de mudança de estado, que licencia (1), é um tipo de construção de estrutura argumental caracterizada por [SN[V[SAdj]]]. Como podemos perceber, essa construção apresenta uma predicação adjetival que é expressa pelo predicado verbo-nominal. Segundo Cunha e Cintra (2013), o predicado verbo-nominal em português possui dois núcleos significativos, um verbo e um predicativo.

Com uma abordagem gerativista, Foltran (1999) apresenta essas construções como predicações secundárias. Ela faz uma divisão entre predicação secundária e mini-oração, sendo que a predicação secundária pode ser orientada para o objeto, para o sujeito, ou ser caracterizada como resultativa. Segundo Foltran, sentenças como (4) são exemplos de predicados secundários orientados para o objeto, enquanto sentenças como (5) são exemplos de predicados secundários orientados para o sujeito. Já

casos como (6) exemplificam predicados secundários resultativos, sendo os casos semelhantes a (7) exemplos de mini-oração.¹⁵

(4) João comprou o carro quebrado

(5) Pedro chegou cansado

(6) Ela cortou o cabelo curto

(7) Eu considero Maria bonita

O que caracteriza, de acordo com Foltran (1999), a predicação secundária é o recebimento de um papel temático do sujeito ou objeto por um outro núcleo lexical, assim temos duas unidades predicadoras. Em (4) e (5), *carro* e *Pedro* são duplamente tematizados, *carro* pelo verbo comprar e pelo adjetivo quebrado e *Pedro* pelo verbo chegar e pelo adjetivo cansado. No caso das mini-orações, o segundo argumento do verbo é expresso por uma oração, como em (7), *Maria bonita*, em que essa predicação não apresenta elemento verbal com flexão de tempo. Por isso, é chamada de mini-oração, pois é o adjetivo o predicator do sintagma nominal. Já a predicação resultativa aplica-se no estado final do objeto direto, apresentando uma causatividade a partir de uma ação verbal, como em (6).

Como podemos observar, a análise realizada por Foltran apresenta um caráter mais composicional e lexicalista, uma vez que a construção de Mudança de Estado é apresentada apenas como resultado de uma dupla tematização de dois predicados.

Já Castilho (2010) afirma que predicados verbo-nominais são minissentenças adjetivais em que temos uma única estrutura desempenhando duas funções, o que evidencia o multifuncionalismo das estruturas sintáticas. No exemplo (8), retirado de Castilho (2010), o autor afirma que *doente* é uma minissentença adjetival que se encontra separada do sintagma verbal *encontraram o povo*, havendo, assim, uma fronteira sintática entre a sentença simples e a minissentença adjetival.

(8) Os pesquisadores encontraram o povo doente.

As descrições realizadas por Castilho e Foltran demonstram casos em que a construção de mudança de estado apresenta a seguinte configuração [SN [V [SN[SAdj]]]], havendo, assim, a existência de outro sintagma nominal entre o verbo e sintagma adjetival como em (4), (6), (7) e (8).

¹⁵ Exemplos retirados de Foltran (1999).

Porém, nesse trabalho para a modelagem de construções, trataremos apenas das construções como em (5).

Nosso objetivo aqui é, primeiramente, demonstrar como uma abordagem construcional pode ser mais eficaz para tratar de estruturas como essas, pois, como afirma Goldberg (1995), estruturas argumentais são instâncias de construções que existem independentemente de verbos – ou outros itens lexicais predicadores, como adjetivos – e evocam frames genéricos das ações humanas básicas, possuindo, assim, significado por si mesmas, o que nos permite diferenciar uma construção de outra. Em segundo lugar, pretendemos ainda demonstrar a possibilidade de modelar tal construção computacionalmente utilizando a infraestrutura disponibilizada pela FrameNet Brasil.

No que concerne ao primeiro objetivo, a AVM apresentada na Figura 1 fornece suporte para o argumento de que uma análise construcional de sentenças como (1) e (5) é mais econômica, além de dar conta, via unificação, de uma variedade maior de possibilidades de preenchimento dos *slots* do verbo e do adjetivo. Considerem-se os exemplos (9-13), extraídos do Corpus do Português – Web/dialetos.¹⁶

- (9) Um belo dia, **Jack amanheceu morto**.
- (10) **João acabou condenado** apenas por homicídio culposo, sem a intenção de matar.
- (11) No último show, em Belgrado, na Sérvia, mês passado, **Amy apareceu bêbada**, esqueceu letras e saiu do palco antes da hora.
- (12) **O Japão acordou feliz** e muito entusiasmado por organizar a maior competição esportiva do mundo.

Apesar de os itens lexicais que ocupam os *slots* do verbo e do adjetivo na construção serem distintos daqueles presentes em (1), guardam algo em comum com eles. Começando pelos verbos, assim como ocorre com *amanhecer*, os verbos *acabar*, *aparecer* e *acordar* não preveem, em sua especificação de valência (val), um argumento interno, pedindo apenas um sujeito. Quanto aos adjetivos, assim como *doente*, os adjetivos *morto*, *condenado*, *bêbado* e *feliz* indicam estados. Nesse sentido, a AVM da Figura 1, que representa o construto em (1), pode ser modificada para representar a construção de Mudança de estado, conforme a Figura 4.

¹⁶ Disponível em <https://www.corpusdoportugues.org/web-dial/>

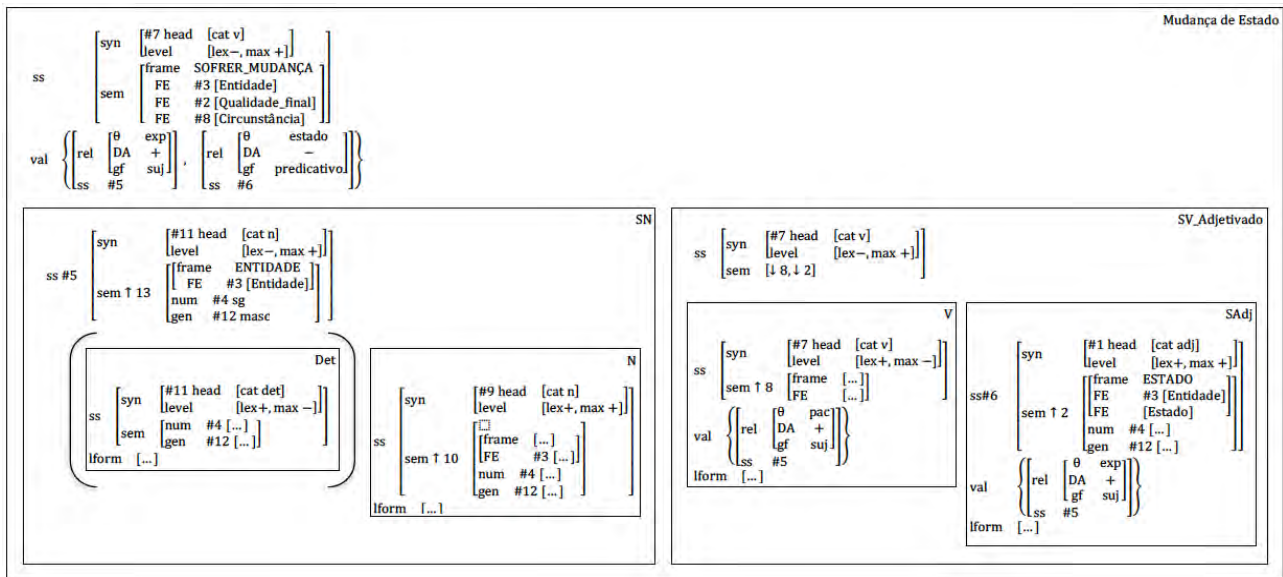


Figura 4: AVM da Construção de Mudança de Estado

Observe-se que valores de diversos atributos encontram-se indefinidos na AVM da construção – [...] – e que o *slot* Det passa a ser opcional. Porém, os índices de unificação se mantêm, de modo a marcar a necessidade, por exemplo, de que o sujeito da construção seja unificado à Entidade a quem o Estado se aplica, ou de que o frame evocado pelo verbo seja unificado a uma determinada Circunstância da mudança de estado, o que é marcado pelo índice #8, que unifica o FE do frame de Sofrer_mudança, ao frame evocado pela LU que preencher o *slot* V.

Já no concernente ao segundo objetivo, passaremos agora a demonstrar de que modo a AVM da Figura 4 pode ser traduzida em uma representação computacional. Dados os princípios e ferramentas de modelagem descritos na abertura da seção 3, a construção de Mudança de estado é definida no Constructicon da FrameNet Brasil conforme a Figura 5.

Mudança de Estado [state_change]

Definição	
Tipo de construção	SN_V_SAdj. Essa construção é composta por um Sintagma Nominal Sujeito e um Sintagma Verbal Adjetival Predicado_Verbo_Nominal . A construção indica uma mudança de estado no Sujeito.
Exemplo(s)	
Elementos da Construção	
Predicado_Verbo_Nominal [predicado verbo nominal]	Predicado composto obrigatoriamente por um Sintagma_verbal_adjetival.
Sujeito [state_change_sujeito]	O sujeito é um Sintagma_Nominal
Relações	
Evoca	Sofrer_mudança
Herda de	Sujeito_predicado

Figura 5: Construção de Mudança de estado no Constructicon da FrameNet Brasil

Como se pode notar pelas descrições, os Elementos da Construção (CEs) são licenciados pelas construções de Sintagma Nominal e Sintagma Verbal Adjetival, que correspondem ao Sujeito e ao Predicado Verbo-Nominal, respectivamente. Para garantir que essa informação conste da representação computacional dessa construção, a restrição CE>Cxn (cxn) é aplicada. Ademais, faz-se necessário indicar que o frame evocado pelo Adjetivo deve evocar algum Estado. Assim, a restrição CE>Frame_family (fam) é aplicada ao núcleo Adjetivo do CE Predicado_verbo_nominal. Por fim, faz-se necessário restringir os verbos que ocorrem no *slot* V àqueles que são admitidos pela construção. Dois tipos de restrição podem ser utilizadas para tanto: a restrição CE>Frame, para os casos em que todos os verbos que evocam um dado frame podem ocupar o *slot*; ou a restrição CE>LU, quando se quiser definir verbos específicos, com sentidos específicos, a ocupar esta posição. A aplicação de tal restrição demandaria um estudo de *corpus* que não está no escopo deste artigo, assim, a título de demonstração, aplicamos a restrição CE>Frame (evk) ao V, para indicar que quaisquer verbos que evoquem o frame de Acordar – *acordar, despertar* – podem preencher essa posição. O resultado da aplicação das restrições é mostrado na Figura 6.

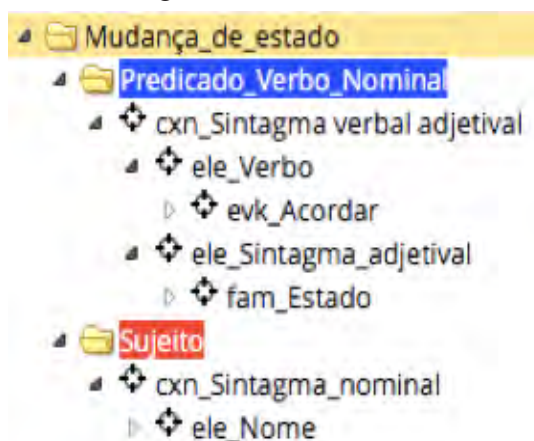


Figura 6: Restrições aplicadas aos CEs da construção de Mudança de estado

Completando a modelagem, as relações de Herança e Evocação, respectivamente, incluem a construção de mudança de estado na rede de construções de Sujeito-Predicado do PB e mapeiam seu aporte semântico para um frame já modelado pela FrameNet, no caso, o frame de Sofrer_mudança. Dessa forma, percebemos que a construção de Mudança de estado herda da construção mais genérica Sujeito-Predicado, sendo que, naquela, o CE Predicado é mais específico do que aquele da construção mãe. Demonstra-se, também, a unificação do polo formal com a contraparte semântica da construção, pela evocação do frame de Sofrer_mudança pela construção, através do mapeamento dos CEs Sujeito e Predicado_verbo_nominal aos FEs Entidade e Qualidade_final, respectivamente.

Vejamos agora como esses princípios de modelagem podem ser aplicados a outras construções

3.2 As Construções de Objeto Interdito, Intransitiva e Transitiva Direta Ativa

A Construção de Objeto Interdito (Bronzato, 2009), se relaciona a contextos de destransitivização do predicador verbal, uma vez que ela omite um complemento tipicamente associado ao rompimento das regras de conduta e a tabus sociais. São licenciadas por essa construção sentenças como aquelas em (13-15), extraídas do Corpus do Português.

- (13) **Uma senhora que bebe** tanto assim, não é uma senhora.
- (14) Por que **você bebe** tanto, mamãe, você fica cheirando mal.
- (15) **O jovem bebe** para se desinibir e participar dos seus grupos.

As três ocorrências acima apresentam o verbo *beber* relacionado ao rompimento de uma conduta social, associada ao consumo de bebidas alcoólicas, geralmente avaliada de forma negativa pelas pessoas. Segundo Bronzato (2009, p. 78), “as sutilezas semânticas, resultantes da destransitivização de predicadores verbais, são frutos de um processo de mesclagem entre o item lexical e a construção sintática”. Desse modo, a Construção de Objeto Interdito atua sobre os frames evocados pelas LUs verbais presentes nas sentenças licenciadas por ela, no sentido de convidar o interlocutor a uma inferência de que o complemento interdito não é costumeiramente bem-visto.

A construção de Objeto Interdito herda da construção Transitiva Direta Ativa, pois a primeira apresenta a valência básica do verbo alterada, ou seja, verbos como *beber*, *cheirar*, *fumar* são considerados, tradicionalmente, conforme Castilho (2010), biargumentais, isto é, apresentam um argumento externo e um argumento interno. Dessa forma, diferentemente da construção Transitiva Direta Ativa, o objeto na construção de Objeto Interdito aparece omitido, uma vez que, conforme Bronzato (2009, p. 78), “o apagamento sintático do participante interdito explica-se pela atuação de uma estratégia de proteção de face, representando, pois, uma ação de *polidez positiva*”.

Nesse sentido, a construção evoca o frame de Avaliação_de_moralidade, segundo o qual um Avaliado (o Sujeito) é julgado quanto a um Comportamento (o Predicado). Em termos de constituição, essa construção é aparentemente similar à construção Intransitiva, uma vez que não apresenta complemento interno ao verbo. Por outro lado, tem o *slot* V preenchido por verbos que preveem, em sua valência, a existência de tal complemento, que foi interdito pela construção.

Assim, o CE Sujeito é licenciado pela construção Sintagma_nominal, enquanto o CE Predicado é, portanto, licenciado pela construção de Sintagma_verbal_sem_complemento, diferentemente do que ocorre com a construção Transitiva Direta Ativa, que tem Predicado licenciado pela construção de Sintagma_verbal_com_complemento.

Até este ponto, em termos de restrições, as construções de Objeto Interdito e Intransitiva são idênticas, o que representa um problema para o modelo. Tal problema é resolvido justamente modelando-se o fato de que, apesar de nuclear um SV_sem_complemento, o V da construção de Objeto Interdito é, normalmente, biargumental. Na FrameNet, verbos biargumentais evocam frames da família de Ação_transitiva, enquanto verbos monoargumentais, tipicamente presentes na construção Intransitiva, evocam frames herdeiros de Ato_intencional. Assim, as restrições CE>Construction e CE>Frame_family, juntamente com a relação de Evocação, são capazes de diferenciar as construções Transitiva Direta Ativa, Intransitiva e de Objeto Interdito entre si, conforme demonstram as Figuras 7 (a, b e c), 8 (a, b e c) e 9 (a, b e c).

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
Herda por	Objeto_interdito
Herda de	Sujeito_predicado

Figura 7(a): Construção Transitiva Direta Aitva

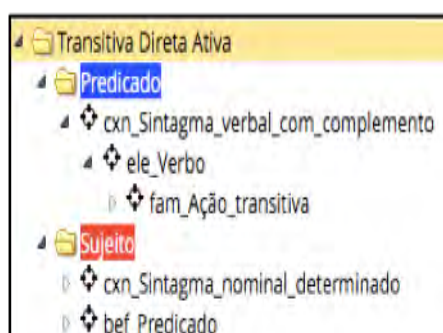


Figura 7(b): Restrições aplicadas aos CEs da Construção Transitiva Direta Ativa

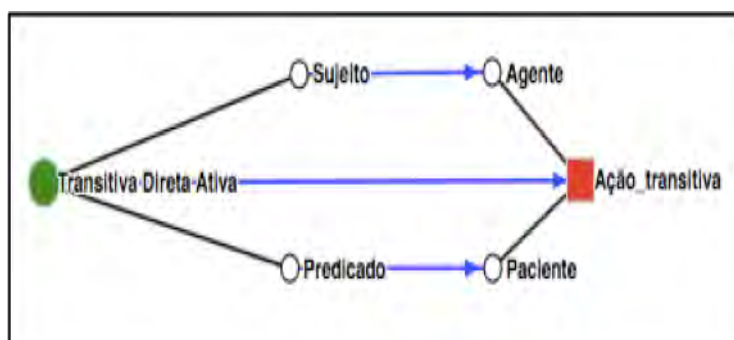


Figura 7(c): Frame evocado pela Construção Transitiva Direta Ativa

Intransitiva [Intransitive]

Definição
 Tipo de construção [SN[V]]. Essa construção exibe um argumento externo **Sujeito** e um **Predicado** sem complemento interno.

Exemplo(s)

Elementos da Construção
Predicado [Predicate] O **Predicado** é um Sintagma verbal sem complemento.
Sujeito [Subject] O **Sujeito** é um Sintagma nominal.

Relações
 Evoca Agir_intencionalmente
 Herda de Sujeito_predicado

Figura 8(a): Construção Intransitiva

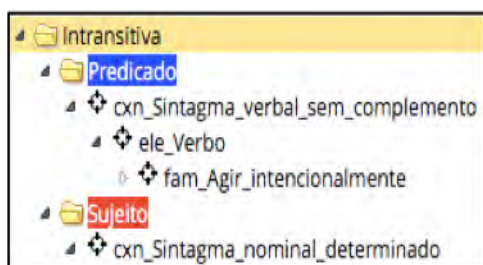


Figura 8(b): Restrições aplicadas aos CEs da Construção Intransitiva

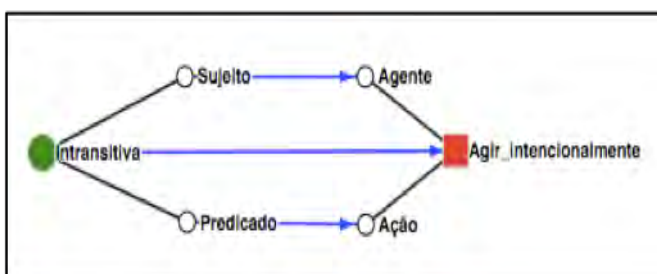


Figura 8(c): Frame evocado pela Construção Intransitiva

Objeto_interdito [object_interrupted]

Definição
 Tipo de construção [[SN] [V[Ø]]]. Essa construção exibe a valência alterada dos verbos, apresentando interdição, ou seja, o objeto está implícito na sentença.

Exemplo(s)

Elementos da Construção
Predicado [Predicate] O predicado é composto por um verbo cujo objeto direto e deliberadamente interditado.
Sujeito [Subject] O sujeito é um SN.

Relações
 Evoca Avaliação_de_moralidade
 Herda de Transitiva Direta Ativa

Figura 9(a): Construção de Objeto Interdito

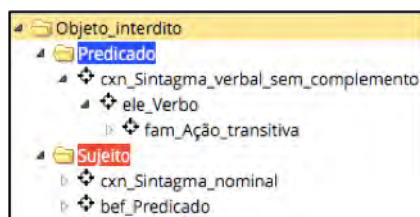


Figura 9(b): Restrições aplicadas aos CEs da Construção de Objeto Interdito

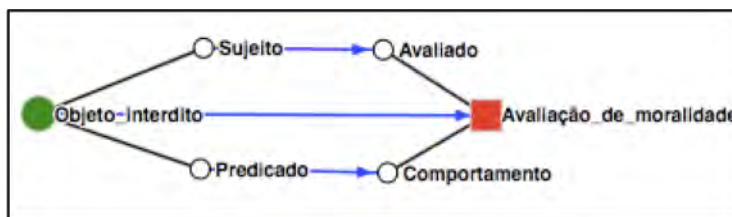


Figura 9(c): Frame evocado pela Construção de Objeto Interdito

Os conjuntos de Figuras 7, 8 e 9 demonstram de que maneira a infraestrutura baseada em restrições e relações do Constructicon da FrameNet Brasil possibilita a representação computacional detalhada das construções em questão. Poder-se-ia questionar, entretanto, que a construção de Objeto Interdito evoca também o frame de Ação_transitiva e que tal informação não consta da Figura 9(c). Porém, há que se considerar que a relação de herança entre esta construção e a Transitiva Direta Ativa registra que tudo o que é válido para a construção mãe também o é para a construção filha, incluindo a evocação de frames. Nesse sentido, o sistema como um todo registra que a construção de Objeto Interdito evoca, via herança, o frame de Ação_transitiva.

3.3 As Construções de Argumento Cindido e Ergativa

A construção Ergativa é um tipo de construção de estrutura argumental que apresenta a seguinte configuração sintática [SN [SV]] em que um sujeito sofre ação de um evento expresso pelo verbo. Diferentemente da construção Intransitiva, que também apresenta uma estrutura monoargumental, o sujeito da Ergativa exhibe comportamento típico do objeto direto dos verbos transitivos. Dessa forma, o argumento interno que é demovido para a posição do sujeito é afetado por uma causa potencialmente externa (PERINI, 2010).

A construção Ergativa licencia construtos como (16) e (17) e, como podemos observar, o sujeito dessa construção não pratica nenhuma ação, sendo paciente.

(16) O vaso quebrou

(17) O lençol rasgou

No processo de representação dessa construção no Constructicon, conforme Figuras 10(a) e 10(b), foi necessário considerar esses aspectos sintáticos-semânticos, uma vez que a construção Ergativa também é herdeira da construção Sujeito-Predicado e evoca o frame Ser_afetado, indicando

que, enquanto o CE Sujeito corresponde a entidade que sofre a ação, o CE Predicado corresponde ao evento com apenas um núcleo verbal, que não é manifesto sintaticamente na forma de um novo argumento.

Ergativa [cxn_ergative]

Definição
Tipo de construção SN_V. Essa construção exibe um argumento interno Sujeito e um Predicado .
Exemplo(s)
Elementos da Construção
Predicado [ce_predicate] O Predicado é composto por um SV.
Sujeito [ce_subject] O Sujeito é um SN.
Relações
Evoca Ser_afetado
Herda de Sujeito_predicado

Figura 10(a): Construção Ergativa



Figura 10(b): Frame evocado pela Construção Ergativa

Já as construções de Argumento Cindido (CACs) são definidas por Sampaio (2010) como construções que apresentam dois argumentos sintáticos e apenas um argumento semântico com função de paciente. Apesar de essas construções apresentarem uma sintaxe semelhante à da construção Transitiva direta ativa, não existe nenhuma ação praticada pelo sujeito que afete um paciente. O que ocorre, na verdade, é que uma mesma entidade se comporta como sujeito e objeto, ambos pacientes, como ilustram o exemplos (18) e (19).

(18) O celular quebrou a tela

(19) A calça rasgou o bolso.

Sendo assim, o que caracteriza as CACs, além da sua estrutura biargumental, e a expressão de um evento que afeta um paciente, é a relação meronímica entre seus argumentos.

Sampaio (2010) afirma que essa relação é a principal característica da CAC-artefato¹⁷, uma vez

¹⁷ Sampaio (2010) postulou uma família de Construções de Argumento Cindido, porém, neste trabalho, apresentamos apenas a representação da CAC-artefato.

que o sujeito é um artefato e seu objeto, uma parte específica desse artefato: a tela consiste em uma parte do celular e o bolso uma parte da calça.

Argumento cindido [cxn_split_object]

Definição
Tipo de construção SN_V_SN. Essa construção exibe um argumento externo Sujeito e um Predicado . O argumento interno é um objeto direto. A construção apresenta uma relação PARTE/TUDO entre os argumentos.
Exemplo(s)
Elementos da Construção
Predicado [ce_predicate_split_object] O Predicado é composto de um SV e um SN objeto direto.
Sujeito [ce_subject_split_object] O sujeito é um SN.
Relações
Evoca Parte_todo, Ser_afetado
Herda de Sujeito_predicado

Figura 11(a): Construção de Argumento Cindido

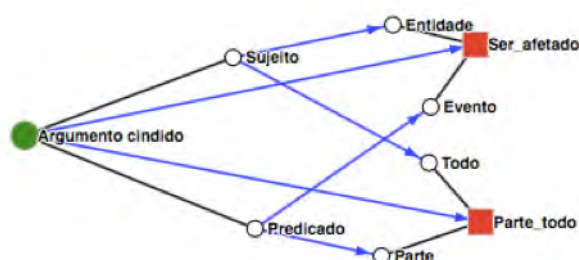


Figura 11(b): Frames evocados pela Construção de Argumento Cindido

Como ilustram as Figuras 11(a) e 11(b), a representação da construção de Argumento Cindido é muito semelhante à representação da construção Transitiva, do ponto de vista formal. A diferença reside justamente na semântica dessas construções: enquanto na CAC os frames Ser_afetado e Parte_Todo estão unificados aos CEs Sujeito e Predicado, através dos FEs Entidade e Evento, e Parte e Todo, a construção Transitiva indica que há um paciente que sofre a ação transitiva de um agente.

Quando comparamos as construções Ergativa, Argumento Cindido e Transitiva Direta Ativa, percebemos a necessidade de um sistema de representação que seja eficiente para tratar de aspectos que vão além da estrutura formal. Se, por um lado, a CAC apresenta a mesma configuração sintática da Transitiva e, por outro lado, apresenta uma contraparte semântica semelhante à da Ergativa, apenas um modelo de representação que se baseie em pareamentos de forma e sentido – a definição clássica de construção (cf. Goldberg, 1995) – permite que se produza um representação eficaz dos fenômenos linguísticos.

3.4 As Construções Predicativas

Completando o conjunto de construções de estrutura argumental do PB modeladas até então no Constructicon da FrameNet Brasil, temos as construções Predicativas, chamadas por Perini (2010) de Estativas, por conta do fato de o *slot* V nessas construções ser preenchido por verbos de estado, tais como *ser*, *estar*, *continuar*, *ficar* e *parecer*. Segundo este autor (2010, p. 101-102), há dois subtipos de construções Estativas: a Estativa propriamente dita, em que uma Qualidade é atribuída a um Qualificando, e a Estativa de Lugar, em que um Lugar é atribuído a um Localizando. As sentenças (20-23), extraídas de Perini (2010), exemplificam essas construções.

- (20) A Letícia é muito inteligente.
- (21) A Letícia é um gênio.
- (22) A toalha está aqui.
- (23) Poços de Caldas é no Sul de Minas.

Analisando-se (20) e (21), exemplos da construção Estativa, nota-se que o *slot* do predicativo pode ser preenchido tanto por um núcleo adjetival, quanto por um núcleo nominal. Já em (22) e (23), esse mesmo *slot* pode ser preenchido tanto por núcleos adverbiais, quanto por Sintagmas Preposicionais.

A análise de Perini (2010) é adaptada à modelagem dessas construções no Constructicon da FrameNet Brasil, as quais, porém, são nomeadas como construção Predicativa Nominal e construção Predicativa Locativa. Tal renomeação decorre da constatação de que o esquema $[SN[V_{cop}[SN|SAdj]]]$, que caracteriza a construção Estativa de Perini (2010) se presta a relacionar uma entidade – o Sujeito – tanto a um estado, quanto a um atributo. Assim, há dois subtipos de construção Predicativa Nominal: a Atributiva, que licencia sentenças como (20), e a Estativa, que licencia casos como (24).

- (24) O peixe está doente.

Dada esta divisão, enquanto a construção Predicativa Nominal Atributiva evoca o frame de Atributos, a Estativa evoca o de Estado. Além disso, restrições se aplicarão sobre os verbos de ligação, ou cópulas, que podem figurar em cada subtipo. Enquanto a versão Atributiva aceita todos os verbos de ligação listados por Perini (2010), a Estativa não aceita o verbo *ser*, uma vez que a noção de permanência advinda desse verbo não condiz com a noção de estado.

As Figuras 12, 13 e 14 mostram a modelagem das construções predicativas nominais, enquanto

a Figura 15 mostra o grafo que representa as relações de herança e evocação propostas para as construções Predicativas nominais.

Predicativa_nominal_atributiva [Attributive_nominal_predicate]

Definição	Tipo de construção predicativa em que um SAdj ou SN codificam um atributo do Sujeito .
Exemplo(s)	
Elementos da Construção	<p>Predicado [Predicate] O Predicado é composto por um SAdj ou SN que codificam um atributo do Sujeito.</p> <p>Sujeito [Subject] O Sujeito é um SN.</p>
Relações	<p>Evoca Atributos</p> <p>Herda de Predicativa_nominal</p>

- ▲ Predicativa_nominal_atributiva
 - ▲ Predicado
 - ◆ cxn_Sintagma_verbal_cópula
 - ▷ ele_Cópula
 - ◆ ele_Predicativo_do_sujeito
 - ▷ fam_Atributos
 - ▲ Sujeito
 - ▷ cxn_Sintagma_nominal

Figura 13: Construção Predicativa nominal atributiva e restrições aplicadas a seus CEs

Predicativa_nominal_estativa [Stative_nominal_predicate]

Definição	Tipo de construção [SN[Vcop[SAdj SN]]] em que um adjetivo ou nome predicativos atribuem um estado a um Sujeito .
Exemplo(s)	
Elementos da Construção	<p>Predicado [Predicate] O Predicado é composto por um SAdj] ou SN predicativos.</p> <p>Sujeito [Subject] O Sujeito é um SN.</p>
Relações	<p>Evoca Estado</p> <p>Herda de Predicativa_nominal</p>

- ▲ Predicativa_nominal_estativa
 - ▲ Predicado
 - ◆ cxn_Sintagma_verbal_cópula
 - ▷ ele_Cópula
 - ◆ ele_Predicativo_do_sujeito
 - ▷ fam_Estado
 - ▲ Sujeito
 - ▷ cxn_Sintagma_nominal

Figura 14: Construção nominal estativa e restrições aplicadas a seus CEs

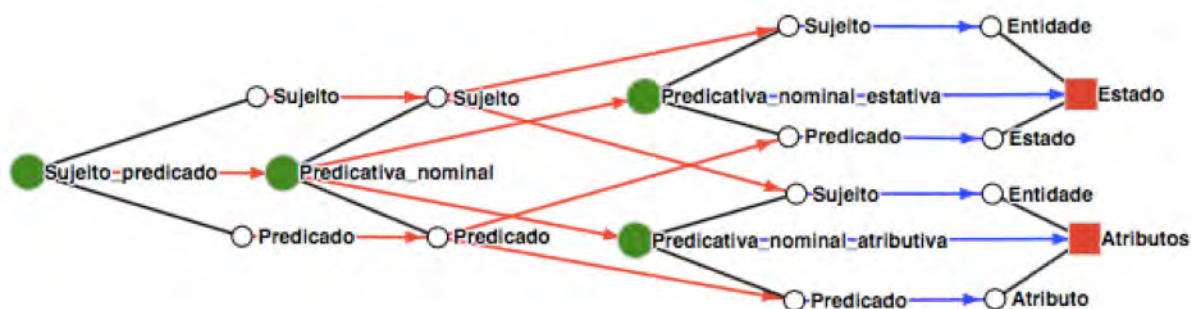


Figura 15: Relações de herança e evocação modeladas para as construções Predicativas nominais

Já a construção Predicativa Locativa segue o esquema sintático [SN [V_{cop} [SAdv|SP]]] e evoca o frame Ser_localizado. Tal construção licencia sentenças como (22) e (23) e foi modelada conforme a Figura 16.

Predicativa_locativa [Locative_predicate]	
Definição	Tipo de construção em que o núcleo é um advérbio ou nome Predicativo que atribui uma localização a um Sujeito .
Exemplo(s)	
Elementos da Construção	<p>Predicado [Predicate] O predicado é composto por uma cópula e um SAdv ou SP.</p> <p>Sujeito [Subject] O Sujeito é um SN.</p>
Relações	<p>Evoca Ser_localizado</p> <p>Herda de Sujeito_predicado</p>

Figura 16: Construção Predicativa locativa

A informação sobre a constituição do CE Predicado é modelada aplicando-se restrições tanto ao CE em si, quanto à construção que o licencia. Assim, utilizando-se a restrição CE>Construction, modela-se que o CE Predicado é licenciado pela construção de Sintagma_verbal_copular_locativo – cf. Figura 17 –, a qual, por sua vez, é composta pelos CEs Cópula e Locativo. Em seguida, aplicam-se ao CE Cópula as restrições CE>Construction, para indicar que ela é licenciada pela construção de Verbo; CE_before, para indicar que a Cópula vem antes do Predicado, e CE>Lexeme, para indicar que apenas os verbos *ser*, *estar*, *ficar* e *continuar* podem ser usados nesse *slot*, excluindo-se o verbo *parecer*, conforme aponta Perini (2010) – cf. Figura 18.

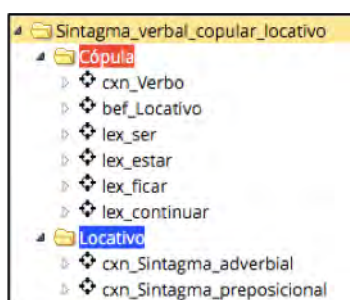


Figura 17: Restrições aplicadas aos CEs da Construção Predicativa locativa

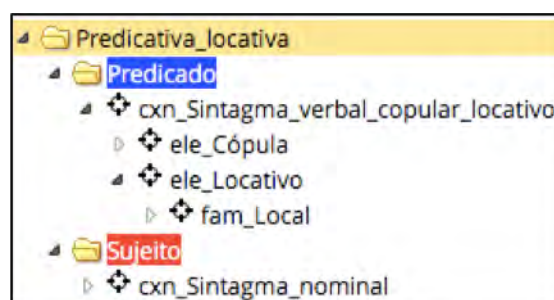


Figura 18: Restrições aplicadas aos CEs da Construção de Sintagma verbal copular locativo

Uma vez modeladas as construções, em termos de suas restrições e relações de herança e

evocação de frames, é possível propor uma rede das construções de Sujeito-Predicado do PB, conforme a Figura 19.

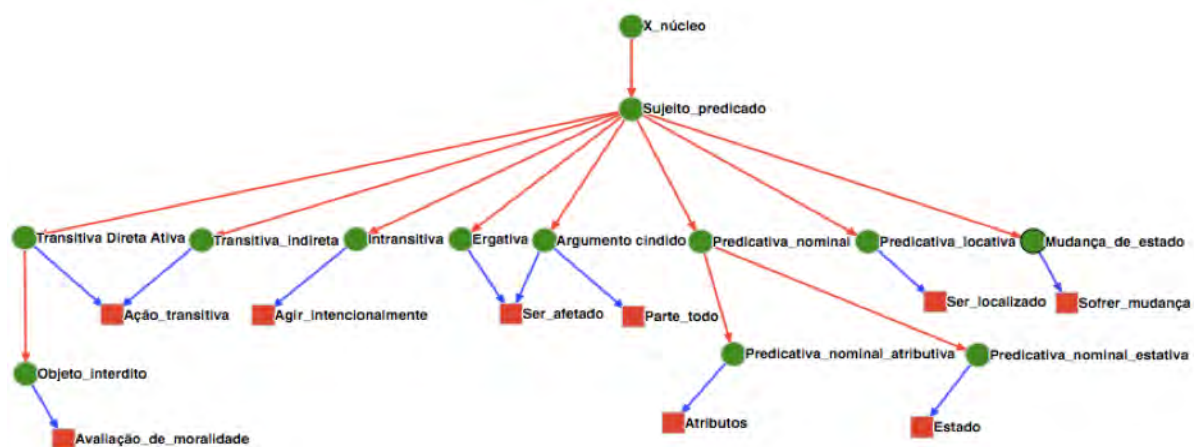


Figura 19: Rede de construções de Sujeito-Predicado do PB

Considerações Finais: Limites à Representação Computacional do Conhecimento

As pesquisas em Compreensão da Língua Natural, como a que vem sendo realizada no âmbito da FrameNet Brasil, apontam para a necessidade de representação de pelo menos dois tipos de conhecimento (Ovchinnikova, 2012): o conhecimento linguístico e o conhecimento “de mundo” (*world knowledge*).

O conhecimento “de mundo” está geralmente associado a um domínio específico e, via de regra, é representado de maneira formal. A formalidade visa eliminar ambiguidades, restringir interpretações, facilitar a comunicação e, em alguns casos, possibilitar uma representação computacional do conhecimento. Para isso, são definidas linguagens formais, com regras estritas. Exemplos destas representações podem ser encontrados nas redes semânticas, nos thesaurus, nas ontologias, nos modelos em grafo, nos modelos conexionistas, nos modelos baseados em lógicas formais, entre vários outros.

O conhecimento linguístico, por outro lado, diz respeito a como o conhecimento “de mundo” é expresso usando uma língua natural. Envolve diversas e variadas dimensões, que tratam das questões fonéticas, morfológicas, sintáticas, semânticas e pragmáticas. Em especial, aborda a complexa questão do significado e de como o significado pode ser representado. As expressões linguísticas – motivadas pela cognição humana – são inerentemente flexíveis, ambíguas, dependentes do contexto, movidas pela criatividade e pouco aderentes a regras (Fauconnier & Turner, 2002).

É natural que, em busca de uma representação computacional para o conhecimento linguístico, a relação da linguagem com outros sistemas cognitivos tenha sido negligenciada. A linguagem foi considerada autônoma e objetiva, e sua representação reduzida a regras lógico-formais. Os estudos na Linguística Cognitiva, no entanto, já mostraram a insuficiência do significante para tratar da dimensão da significação linguística (Salomão, 1999; Fauconnier & Turner, 2002).

O problema não é representar o conhecimento através do uso sistemático da forma. O problema é considerar que essas representações abarcam todas as dimensões do significado. A formalização do conhecimento linguístico é necessária no âmbito computacional, mas não se deve ignorar a limitação que esta formalização impõe. Confiar que as formas são capazes de processar a essência dos sistemas conceptuais de conhecimento é dizer que a capacidade humana de construir significados pode ser manipulada totalmente. No entanto, como afirmam Fauconnier e Turner (2002), as expressões linguísticas apontam para o significado, mas não o portam.

Esses limites à representação computacional do conhecimento podem ser amenizados por teorias linguísticas como a Semântica de Frames e algumas abordagens da Gramática das Construções, que buscam ser cognitivamente plausíveis. Como foi visto, a FrameNet e o Constructicon aliam o registro dos fenômenos linguísticos ao registro do significado associado a esses fenômenos. A rede de frames, elementos de frame, unidades lexicais e construções, representando simultaneamente o conhecimento linguístico e o conhecimento “de mundo”, constitui uma ferramenta flexível e abrangente para o estudo e a compreensão da língua sob o ponto de vista computacional.

As construções que foram aqui modeladas explicitam a necessidade de considerar tanto aspectos semânticos como aspectos sintáticos para uma representação satisfatória do conhecimento linguístico. Para isso, de um lado, as restrições sintáticas foram tratadas em termos de sintagmas e categorias gramaticais como uma forma de mostrar as generalizações presentes em cada construção. Por outro lado, cada restrição foi unificada a um determinado componente semântico do frame evocado pela construção.

A proposta é abarcar no Constructicon diferentes construções de estrutura argumental com o intuito de constituir uma gramática do PB que inclua diferentes padrões construcionais, já que esses manifestam frames genéricos das ações humanas básicas, que já estão disponíveis na base da FrameNet Brasil.

REFERÊNCIAS

Allen, J. (1995). *Natural Language Understanding*. New York: Pearson.

Almeida, V. G. (2016). *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.

Bronzato, L. H. (2009). O enquadre gramatical da interdição ou “a bom entendedor meia palavra basta”. In: Miranda, N. S. & Salomão, M. M. M. (orgs.). *Construções do Português do Brasil: da gramática ao discurso*. Belo Horizonte: UFMG, p. 76-97.

Bryant, J. E. (2008). *Best-fit constructional analysis*. Tese de Doutorado em Ciência da Computação, University of California, Berkeley.

Castilho, A. T. (2010). *Nova Gramática do Português Brasileiro*. São Paulo: Editora Contexto.

Chomsky, N. (1993 [1981]). *Lectures on government and binding: The Pisa lectures*. Berlin: Mouton de Gruyter.

Cunha, C. F. & Cintra, L. (2013). *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira.

Dominey, P. F., Meallier, A. L., Pointeau, G., Mirliaz, S. & Finlayson, M. (2017). Dynamic construction grammar and steps towards the narrative construction of meaning. In *The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02*. Palo Alto, CA: AAAI Publications, v.17, p.163-170.

Fauconnier, G. & Turner, M. (2002). *The Way We Think*. New York: Basic Books.

Fillmore, C. J. (1979). Innocence: a second idealization for linguistics. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: BLS, v. 5, p. 63-76.

_____. (1982). Frame semantics. In: Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin, p.111-138.

_____. (1985). Frames and the Semantics of Understanding. *Quaderni di semantica*, 6 (2), p. 222-254.

_____. (2007). Valency Issues in FrameNet. In: Herbst, T. & Götz-Votteler, K. (orgs.). *Valency: Theoretical, Descriptive and Cognitive Issues*. Berlin and New York: Mouton de Gruyter, p. 129–160.

_____. (2008). Border conflicts: FrameNet meets construction grammar. In Bernal, E. & DeCesaris,

J. (eds.). *Proceedings of the XIII EURALEX International Congress*. Barcelona: Universitat Pompeu Fabra, v. 4968, p. 49-68.

_____. (2013). Berkeley Construction Grammar. In: Hoffmann, T. & Trousdale, G. (eds.). *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, p. 111-132.

_____.; Kay, P. & O'Connor, M. (1988). Regularity and idiomacity in grammatical constructions: the case of let alone. *Language*, 64 (3), p. 501-538.

_____.; Petruck, M. R.; Ruppenhofer, J. & Wright, A. (2003). FrameNet in action: The case of attaching. *International journal of lexicography*, 16 (3), p. 297-332.

Foltran, M. J. (1999). As construções de predicação secundária no português do Brasil: aspectos sintáticos e semânticos. Tese de Doutorado em Linguística. Universidade Federal do Paraná, Curitiba.

Fried, M. & Östman, J-O. (2004). Construction Grammar: A thumbnail sketch. In: _____. (orgs.). *Construction Grammar in a Cross-Linguistic Perspective*. Amsterdam: John Benjamins, p. 11-86.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Kay, Paul; Fillmore, Charles J. (1999). Grammatical constructions and linguistic generalizations: The 'What's X doing Y?' Construction. *Language*, 75 (1), p.1-33.

Marques, T. & Beuls, K. (2016). A Construction Grammar Approach for Pronominal Clitics in European Portuguese. In *International Conference on Computational Processing of the Portuguese Language*. Berlim: Springer, p. 239-244.

Lage, L. M. (2018). *Modelagem Linguístico-Computacional das Relações entre Construções e Frames no Constructicon da FrameNet Brasil*. Tese de Doutorado em Linguística. Universidade Federal de Juiz de Fora, Juiz de Fora.

Lyngfelt, B. (no prelo). Introduction: constructicons and constructicography. In: Lyngfelt, B.; Borin, L.; Ohara, K. & Torrent, T. T. *Constructicography: Constructicon Development across Languages*. Amsterdam: John Benjamins.

Matos, E. E.; Torrent, T. T.; Almeida, V. G.; Silva, A. B. L.; Lage, L. M.; Marção, N. D.; Tavares, T. S. (2017). Constructional Analysis Using Constrained Spreading Activation in a FrameNet-Based Structured Connectionist Model In: *The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02*. Palo

- Alto, CA: AAAI Publications, v.17, p.222-229.
- Minsky, M. (1975). A framework for representing knowledge. In: Winston, P. (org.). *The psychology of computer vision*. New York: Mc Graw-Hill.
- Ovchinnikova, E. (2012). *Integration of World Knowledge for Natural Language Understanding*. Paris: Atlantis Press.
- Perini, M. A. (2010). *Gramática do Português Brasileiro*. São Paulo: Parábola.
- Poole, D. L., & Mackworth, A. K. (2010). *Artificial Intelligence: foundations of computational agents*. Cambridge: Cambridge University Press.
- Pustejovsky, J. & Stubbs, A. (2012). *Natural Language Annotation for Machine Learning: A guide to corpus building for applications*. Sebastopol, CA: O'Reilly.
- Sag, I. Sign-Based Construction Grammar: an informal synopsis. In: Boas, H. C. & Sag, I. (orgs.). (2012). *Sign-Based Construction Grammar*. Stanford: CSLI Publications.
- Salomão, M. M. M. (1999). A questão da construção do sentido e a revisão da agenda dos estudos da linguagem. *Veredas*, 3 (1), p. 61-79.
- _____. (2009). FrameNet Brasil: um trabalho em progresso. *Calidoscópio*, 7(3), p. 171-182.
- Sampaio, T. F. (2010). *A Família de Construções de Argumento Cindido no Português do Brasil*. Tese de Doutorado em Linguística. Universidade Federal de Juiz de Fora, Juiz de Fora.
- Silva, A. B. L.; Lage, L. M.; Marcao, N. D.; Tavares, T. S.; Almeida, V. G.; Matos, E. E. & Torrent, T. T. (2017). In: *The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding Technical Report SS-17-02*. Palo Alto, CA: AAAI Publications, v.17, p.193-196.
- Steels, L. (Ed.). (2011). *Design patterns in fluid construction grammar*. Amsterdam: John Benjamins Publishing.
- Torrent, T. T. & Ellsworth, M. (2013). Behind the Labels: criteria for defining analytical categories in FrameNet Brasil. *Veredas*, 17 (1), p. 44-65.
- Torrent, T. T.; Lage, L. M.; Sampaio, T. F.; Tavares, T. S. & Matos, E. E. S. (2014). Revisiting border conflicts between FrameNet and Construction Grammar: Annotation policies for the Brazilian Portuguese Constructicon. *Constructions and Frames*, 6 (1), p. 34-51.