

Construções de Estrutura Argumental no âmbito do Constructicon da FrameNet Brasil: proposta de uma modelagem linguístico-computacional

Vânia Gomes de Almeida¹, Tiago Timponi Torrent¹

¹FrameNet Brasil – Programa de Pós-Graduação em Linguística
Universidade Federal de Juiz de Fora (UFJF)
Rua José Lourenço Kelmer s/nº - Campus Universitário
36036-900 – Juiz de Fora – Minas Gerais – Brasil

vania.almeida2017@letras.ufjf.br, tiago.torrent@ufjf.edu.br

Abstract. *This paper aims to present the modeling of argument structure constructions in the FrameNet Brasil Constructicon. We assume that constructions are basic units of language that constitute correspondences between form and meaning [Goldberg, 1995; Kay & Fillmore 1999]. Therefore, we model the Active Direct Transitive Construction, the Ergative Construction and the Split Argument Construction according to their syntactic-semantic characteristics. The results indicate how constructional modeling can contribute to Natural Language Processing tasks.*

Resumo. *O presente trabalho visa apresentar a modelagem de construções de estrutura argumental no Constructicon da FrameNet Brasil. Assumimos que construções são unidades básicas da língua que se constituem em correspondências entre forma e significado [Goldberg 1995; Kay & Fillmore 1999]. Diante disso, buscou-se modelar as Construções Transitiva Direta Ativa, Ergativa e Argumento Cindido de acordo com suas características sintático-semânticas. Os resultados indicam como a modelagem construcional pode contribuir para tarefas de Processamento de Língua Natural.*

1. Introdução

A FN-Br ocupa-se do desenvolvimento de dois grandes recursos computacionais: um Lexicon e um Constructicon. O Lexicon tem por objetivos: criar uma representação computacional de frames, definidos por seus participantes e instrumentos e conectados entre si via relações entre frames; definir Unidades Lexicais (ULs), pareamentos entre um lema e um significado definido em termos de um frame, e anotar sentenças que exemplifiquem os padrões de valência sintáticos e semânticos em que as ULs ocorrem.

Já o Constructicon visa à criação de um recurso para a descrição das características gramaticais de construções do PB, incorporando descrições interpretáveis computacionalmente para cada construção, oferecendo informações semânticas e especificando as relações entre as construções.

Por essa razão, o Constructicon abarca o conhecimento linguístico que excede a valência simples de palavras simples [Fillmore 2008]. No contexto desses recursos, apresentaremos a modelagem linguístico-computacional de três construções de estrutura

argumental do PB, a saber, Construção Transitiva Direta Ativa, Construção Ergativa e Construção de Argumento Cindido.

2. Lexicon e Constructicon na FrameNet Brasil

Uma vez que partimos do princípio de que construções são unidades básicas da língua constituídas pela correspondência entre forma e significado [Goldberg 1995; Kay & Fillmore 1999], assumimos também a existência de um continuum entre léxico e sintaxe [Fillmore 2008]. Assim, não se trata de haver módulos rigidamente separados, mas, sim, um continuum de construções que partem de elementos mais lexicalmente específicos para padrões mais abstratos. Isso significa dizer que Lexicon e Constructicon não são dois recursos totalmente separados, ambos se conectam na medida em que: (a) itens lexicais são construções e são, portanto, licenciados por construções lexicais; (b) tanto itens lexicais quanto construções não lexicais evocam frames, os quais constituem a categoria analítica primeira da FrameNet. Dessa forma, as duas frentes estão interligadas a fim de proporcionar, tanto no ambiente lexical como no construcional, uma descrição adequada dos fenômenos linguísticos.

Assim, a FN-Br realiza tanto anotações lexicográficas, quanto construcionais. Tais anotações são realizadas em camadas que contribuem com informações sobre as propriedades semânticas e sintáticas dos elementos em análise. A Figura 1 apresenta um exemplo de anotação.

<div> <div>Save</div> <div>Refresh</div> <div>Hide AS</div> <div>Remove AS</div> <div>Label Help</div> </div>									
[13679]	AST_MS_APP	NI	0	c	a	b	o	d	a
CE			S	u	j	e	i	t	o
Ser_afetado.FE			E	n	t	i	d	a	d
CstrPT			S	f	i	n			
Quebrar.quebrar.v									q
FE			S	u	b	p	a	r	t
FE			T	o	d	o			
GF			E	x	t				
PT			N	P					

Figura 1. Anotações lexicográfica e construcional na FrameNet Brasil

A Figura 1 apresenta a sentença *O cabo da panela quebrou* anotada tanto construcionalmente, para a construção Ergativa, quanto lexicograficamente, para a UL verbal *quebrar*. Nesse exemplo, observamos, na primeira camada, os Elementos da Construção (CEs), o *cabo da panela* e *quebrou* anotados respectivamente como Sujeito e Predicado. Na segunda camada, encontramos a relação entre a construção e o *frame* que ela evoca, *Ser_afetado*, através da unificação entre os CEs e o FEs (Elementos de Frame) do mesmo frame, e a última camada anotada corresponde ao Tipo Sintagmático do construto.

Já para a anotação lexicográfica, apresenta-se a UL que evoca o frame *Quebrar*, cujos FEs *Parte* e *Todo* são ambos instanciados no SN *O cabo (Parte) da panela (Todo)*.

A anotação registra, ainda a Função Gramatical (GF) e o Tipo Sintagmático (PT) do material linguístico que instancia os FEs anotados.

Enquanto a anotação lexicográfica produz, no conjunto das sentenças, os padrões de valência de cada UL, o método construcional mapeia uma construção formalmente, unificando-a a um *frame* específico que resulta no licenciamento de um construto. Nesse exemplo, o construto *O cabo da panela quebrou* é licenciado pela construção Ergativa, unificada ao *frame* Ser_afetado. Através dessa tarefa, pode-se modelar adequadamente a continuidade entre o léxico e a gramática para que as informações sintático-semânticas das estruturas linguísticas sejam analisadas a partir das contribuições de cada elemento descrito.

3. A Modelagem de Construções

A fim de realizar um experimento específico de descoberta construcional por máquina, descrito em Almeida (2016), apresentamos a modelagem de três construções de estrutura argumental. A construção Transitiva Direta Ativa, a construção Ergativa e a construção de Argumento Cindido.

As construções analisadas nesse trabalho são herdeiras da construção X_Núcleo, composta por um núcleo que é especificado por um elemento à esquerda, o que significa, de acordo com Kay e Fillmore (1999), que a construção herdeira contém toda a informação da construção herdada e outras informações acrescidas, referentes aos elementos que preenchem a construção.

A partir das descrições realizadas por Ferreira (2009), Castilho (2010), Perini (2010), que apresentaram estudos sobre as Construções Transitiva Direta Ativa e Ergativa, e de Sampaio (2010) para as Construções de Argumento Cindido, nossa proposta é modelar essas construções linguístico-computacionalmente com o intuito de mostrar como essa modelagem pode contribuir para trabalhos na área de Linguística Computacional.

Perini (2010) define construção como uma estrutura gramatical composta por seus constituintes sintáticos e pela relação semântica que cada um deles tem com verbo da oração. Nesse sentido, a construção Transitiva Direta Ativa é definida pela sequência [SN [V SN]] em que o primeiro sintagma nominal designa quem praticou uma ação e o segundo sintagma nominal designa quem sofreu a ação, como em (1). Já a construção Ergativa tem como característica o fato de o sujeito ser paciente, ou seja, não pratica nenhuma ação e sim sofre o efeito do evento expresso pelo verbo. Pode ser definida como a sequência [SN [V]], como em (2). Por fim, conforme Sampaio (2010), as construções de Argumento Cindido (CACs) são construções que apresentam dois argumentos sintáticos e apenas um argumento semântico, com função paciente, juntamente com uma relação Parte_Todo entre eles. Assim, temos um desencontro (*mismatch*) entre estruturas sintáticas e estruturas semânticas como em (3).

- (1) [[A criança_{SN/Agente}] [furou_V [os balões_{SN/Paciente}]]
- (2) [[O pneu_{SN/Paciente}] [furou_V]]
- (3) [[O pedreiro_{SN/Paciente-Todo}] [furou_V [o dedo_{SN/Paciente-Parte}]]

A partir de suas características sintático-semânticas, as relações fundamentais entre três as construções foram modeladas em uma rede integrada. Assim sendo, serão

explicitadas, de um lado, as relações fundamentais para a constituição de um Constructicon como uma rede integrada de construções e frames, quais sejam, a relação de Herança – entre construções – e a de Evocação – entre construções e *frames*. De outro, apresentam-se as restrições, as quais remetem às características sintáticas e semânticas que são específicas de cada construção. Isso implica que um Elemento da Construção pode ser de um determinado tipo semântico, ou seja, deve ser um Elemento de Frame do *frame* evocado pela construção que desempenhe funções sintáticas específicas, e também de um determinado tipo sintático, isto é, tem sua definição ancorada em construções sintagmáticas ou lexicais mais genéricas, também estas definidas no Constructicon.

Como afirma Goldberg (1995), as construções estão normalmente relacionadas entre si, integrando redes construcionais que podem apresentar diferentes tipos de relações. A ideia é que cada construção tem um significado próprio e uma forma esquemática e convencional que é capaz de fornecer as mais abstratas generalizações para o licenciamento de outras construções. Por exemplo, uma Construção Transitiva, que licencia sentenças como *A professora rasgou o papel*, definida formalmente por [SN [V SN]], fornece um padrão formal para uma Construção de Argumento Cindido, que licencia construtos como *A calça rasgou o bolso*, ao mesmo tempo em que compartilha a informação semântica de uma Construção Ergativa, que, por sua vez, licencia *O lençol rasgou*. A Figura 2 mostra as relações entre as três construções, todas herdeiras da Construção Sujeito_Predicado, que, por sua vez, é herdeira da Construção X_Núcleo.

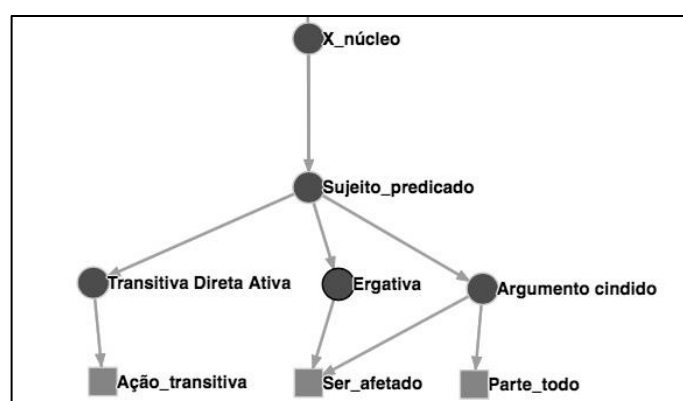


Figura 2. Relações entre frames e construções

As relações de herança são representadas por linhas conectadas aos círculos, que representam as construções. O compartilhamento de estrutura semântica entre a CAC e a Ergativa é representado pela relação de Evocação entre essas construções e o frame *Ser_afetado*, representado por um quadrado.

O princípio que estrutura a relação entre construções é o Princípio da Não Sinonímia, que afirma que, se duas construções são sintaticamente distintas, tais construções devem ser também distintas semântica ou pragmaticamente [Goldberg 1995]. A situação da CAC em relação à Ergativa é exatamente aquela prevista neste princípio. CAC e Ergativa são sintaticamente diferentes, mas semanticamente sinônimas, assim sendo, devem ser pragmaticamente distintas. Os exemplos (1), (2) e (3) mostram que, apesar de a CAC e a Ergativa serem semanticamente sinônimas, há uma diferença entre elas, uma vez que, enquanto na Ergativa encontramos uma estrutura monoargumental, em

que uma entidade é afetada por um evento, a CAC apresenta uma estrutura em que seus dois argumentos possuem uma relação Parte-Todo.

Por outro lado, a CAC é semelhante à Transitiva Direta Ativa sintaticamente, portanto, mesmo sendo semanticamente distintas, deve haver algum grau de motivação entre elas, o que é capturado através do fato de que ambas herdam da construção de Sujeito_predicado.

A Figura 3 apresenta as relações da CAC com as outras construções, como também sua configuração sintático-semântica. Na Construção de Argumento Cindido os CEs, Sujeito e Predicado, que se constituem das construções Sintagma Nominal e Sintagma Verbal com Complemento, correspondem as restrições sintáticas. Os *frames* Ser_afetado e Parte_Todo, que estão representados pelos quadrados, estão relacionados à construção pela relação de Evocação, seus FEs, que estão unificados aos CEs, constituem a Entidade e o Evento no *frame* Ser_afetado, enquanto no *frame* Parte_Todo constituem o Todo e a Parte. A unificação entre os CEs e os FEs representa as restrições semânticas.

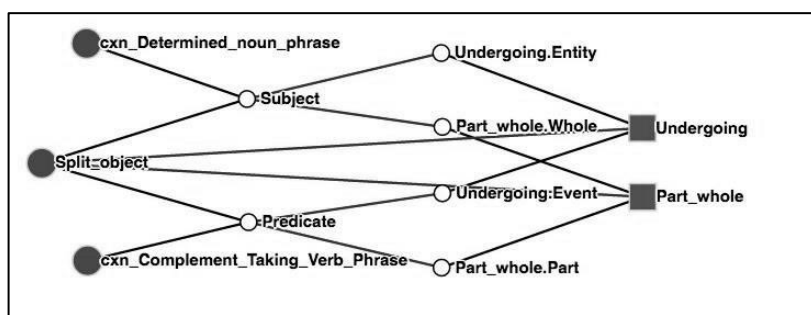


Figura 3. Modelagem da Construção de Argumento Cindido

Na Figura 4, encontramos a modelagem da Construção Transitiva Direta Ativa, que também é herdeira da Construção Sujeito_Predicado e possui como CEs as mesmas construções da Construção de Argumento Cindido. Apesar de a contraparte sintática da CAC ser idêntica à da Transitiva Direta Ativa, notamos, pela relação de Evocação, que a distinção entre elas se encontra na contraparte semântica, já que o frame aqui evocado é de Ação_Transitiva, em que um Agente ou uma Causa afeta um Paciente. Esses FEs, estão unificados aos CEs, que se licenciam pelas construções de Sintagma Nominal e Sintagma Verbal com Complemento.

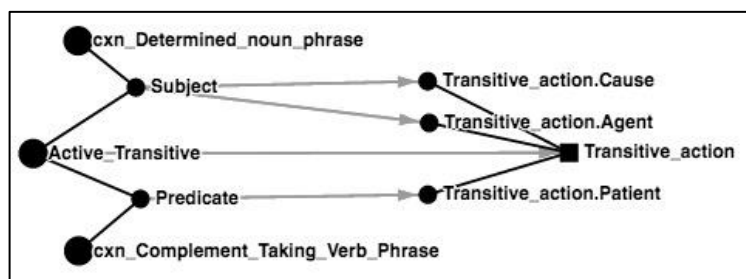


Figura 4. Modelagem da Construção Transitiva Ativa

Por fim, a Construção Ergativa, representada pela Figura 5, também é herdeira da construção Sujeito_Predicado, porém a Construção Ergativa é constituída pelas construções de Sintagma Nominal e Sintagma Verbal sem Complemento, já que a possui uma estrutura monoargumental. Na contraparte semântica, evoca, assim como a CAC, o *frame* Ser_afetado, cujo FE Entidade está unificado ao Sujeito da construção. O Predicado

está unificado ao FE Evento porque, apesar de a construção apresentar dois CEs, a Construção de Sintagma Verbal sem Complemento indica que esse CE é apenas um núcleo verbal, que não é manifesto sintaticamente na forma de um novo argumento.

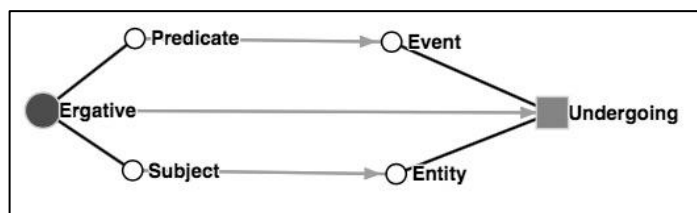


Figura 5: Modelagem da Construção Ergativa

4. Considerações finais

Com a elaboração de uma rede, foi possível demonstrar a unificação entre a contraparte semântica e sintática dessas construções, que foi representada através do mapeamento dos CEs para os FEs dos frames evocados por elas. A modelagem das construções conseguiu mapear que, de um lado, a CAC e a Ergativa são semanticamente sinônimas, mas com uma diferença entre elas, pois, enquanto a Ergativa apresenta uma estrutura monoargumental, em que uma entidade é afetada por um evento, a CAC apresenta uma estrutura de dois argumentos com a mesma informação semântica da ergativa, acrescida por uma relação parte-todo entre eles. Por outro lado, foi possível demonstrar como a CAC é semelhante à Transitiva Direta Ativa sintaticamente, devido a uma motivação que se manifesta pelo fato de ambas as construções herdarem de um tipo específico da construção Sujeito-Predicado, que especifica a transitividade do núcleo do Predicado como direta.

Os resultados obtidos apontam para a importância da interação entre os recursos disponíveis na FN-Br. Conforme apresentado, o Lexicon e o Constructicon precisam funcionar juntos para que a continuidade entre o léxico e gramática seja modelada de forma satisfatória. A perspectiva de tratar os fenômenos semânticos via frames e formalizar o significado das sentenças juntamente com seus aspectos sintáticos fornece à modelagem de construções apresentada nesse trabalho um alinhamento entre um modelo de descrição linguística e a implementação computacional da FN-Br ambos com potencial de contribuir para tarefas de Processamento de Língua Natural.

Referências

- Almeida, V. G. (2016) *Identificação Automática de Construções de Estrutura Argumental*. Dissertação de Mestrado em Linguística, Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil.
- Castilho, A. T. de. (2010) *Nova Gramática do Português Brasileiro*. São Paulo: Contexto.
- Ferreira, M. A. (2009) Construção de Ação Rotineira no Português do Brasil. In Miranda, N. S.; Salomão, M. M. M. *Construções do Português do Brasil: da gramática ao discurso*. Belo Horizonte: UFMG.
- Fillmore, C. J. (2008) Border Conflicts: FrameNet Meets Construction Grammar. In: *EURALEX,13, Barcelona. Anais*. Barcelona: Universitat Barcelona Fabra.

- Goldberg, A. (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Kay, P. & Fillmore, C. J. (1999) Grammatical Constructions and Linguistic Generalizations: the What's X Doing Y Construction. *Language*, vol. 75, nº 1.
- Perini, M. A. (2009) *Gramática Descritiva do Português*. São Paulo: Editora Ática.
- 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, Brasil.