

UM ESTUDO DA UNIDADE LEXICAL CORTAR
SOB À LUZ DA SEMÂNTICA DE FRAMES

Maíra Mendes Magela
mairamagela@hotmail.com

1. *Introdução*

Devido à necessidade de admitir estruturas de conhecimento implícitas (conhecimento de mundo) por meio das quais se operam os processos semânticos de inferenciação, faz-se necessário, desse modo, postular a noção de *frame*. A noção de *frame*, *a priori*, surgiu na Inteligência Artificial e na Sociologia por meio dos estudos desenvolvidos por Marvin Minsky e Erving Goffmann entre as décadas de 70 e 80. Posteriormente esse conceito foi adaptado para o estudo semântico da linguagem por Fillmore em seu artigo (FILLMORE, 1975) dedicado a demonstrar as insatisfações do tratamento componencial da significação lexical, isto é, as insatisfações da descrição do significado lexical de acordo com os termos de uma lista de condições necessárias (componentes ou traços semânticos), cuja conjunção constituiria a condição suficiente para a pertinência de um lexema a uma dada categoria significativa (FILLMORE *apud* SALOMÃO, 2009).

Portanto, caracteriza-se *frame* como uma estrutura conceptual complexa, a qual é organizada de tal forma que para compreender qualquer de suas partes é necessário o conhecimento do todo. Vale destacar que as expressões linguísticas evocam *frames* para serem significativas, apesar de os *frames* não constituírem esquemas conceptuais exclusivamente linguísticos. Assim, resume-se como ponto central desta pesquisa que toda unidade lexical evoca um *frame*, mas, ao fazê-lo, perfila algum elemento desse *frame* de forma particular (SALOMÃO, 2009).

Nesse escopo, esta pesquisa baseou-se tanto teórica quanto metodologicamente no projeto FrameNet Brasil (localizado na Universidade Federal de Juiz de Fora e liderado pela Profa. Dra. Margarida Salomão), o qual, por sua vez, está vinculado ao projeto mãe FrameNet (localizado em Berkeley e liderado pelo prof. Dr. Charles Fillmore). Em suma, o objetivo do projeto é descrever as unidades em português, que tenha como ponto de partida a identificação de esquemas conceptuais evocados por unidades linguísticas por meio de papéis semânticos instanciados no plano sintático dessas expressões.

Desse modo, a partir de evidências extraídas de *corpus*, tal investigação garante a descrição das relações lexicais entre os elementos de *frames* das unidades lexicais, que, geralmente, não são abarcadas pela abordagem tradicional. Além disso, permite a descrição das relações existentes entre os próprios esquemas conceptuais admitidos como componentes de uma rede de *frames*; rede semântica essa que se distingue das ontologias recorrentes nas abordagens computacionais, devido ao seu caráter motivado no uso da língua e que, portanto, abrange os aspectos cognitivos e socioculturais.

Assim, este trabalho de iniciação científica pretendeu descrever a unidade lexical “cortar” como evocadora do *frame* de experiência de dano corporal a fim de contribuir com as pesquisas desenvolvidas no projeto FrameNet Brasil.

Esta proposta de pesquisa se empreendeu, em termos gerais, com o objetivo de analisar e descrever, com base na Semântica de Frames, a unidade lexical “cortar”, a qual evoca o *frame* de experiência de dano corporal, além de anotá-la sistematicamente, por meio de processos automáticos e manuais. De modo mais específico, após descrita e anotada, verificou-se como essa unidade lexical e seu respectivo *frame* contribuiu para a formação de uma rede de *frames*.

2. Metodologia

Para a realização desta pesquisa lexicográfica, fez-se necessária a leitura do material bibliográfico pertinente à temática, baseado, principalmente, nas obras de Fillmore e no *The Book*¹. Paralelo a isso, foi realizado o processo de descrição lexicográfica, o qual constou primeiramente no levantamento das ocorrências da unidade lexical “cortar” a partir de acervos de *corpora* eletrônicos², como AC/DC (em que se encontram os *corpora* ANCIB, ECI-EBR E NILC/São Carlos), disponível no site www.linguateca.pt, e *Sketch Engine* (em que estão os *corpora* Legendas de filmes e NURC-RJ), disponível em www.sketchengine.co.uk.

¹ O *The Book* é a obra que rege o projeto de pesquisa FrameNet. Portanto, nessa obra encontram-se a base tanto metodológica quanto teórica da pesquisa de descrição lexical por meio da Semântica de Frames.

² Os acervos eletrônicos são compostos por *corpora* de variados gêneros textuais, tais como: oral, didático, jurídico, literário, técnico/ científico, jornalístico e mensagem eletrônica.

Em seguida, deu-se início ao procedimento de anotação das ocorrências selecionadas em três camadas, nas quais serão identificados: 1) os elementos do *frame* pertinentes ao *frame* de experiência de dano corporal evocado pela unidade lexical “cortar”, 2) a expressão sintagmática de cada um dos elementos do *frame* em estudo e 3) a função gramatical de cada constituinte sentencial.

Por último, a junção das ocorrências levantadas, descritas sintática e semanticamente, que foram rotuladas a fim de apresentar a descrição da construção da unidade lexical “cortar” por meio de padrões de valências.

3. Procedimentos de Anotação

A pesquisa se fundamentou na definição do *frame* *Experience_bodily_harm*, da língua inglesa, disponível no site do projeto FrameNet mãe¹. Desse modo, fez-se uma tradução para o português da descrição do *frame*, bem como da unidade lexical “cortar”. Assim, o *frame* experiência de dano corporal caracteriza-se por um **experenciador** que se envolve em uma lesão corporal de uma **parte do corpo**. Mas, em alguns casos, nenhuma **parte do corpo** precisa ser indicada. Frequentemente, uma **entidade danificadora** na qual o **experenciador** se fere é mencionada.

Conforme a postulação do *frame* de experiência de dano corporal, os elementos de *frame* se dividem em dois principais grupos: nucleares e não nucleares. Os nucleares são parte do corpo e experienciador, enquanto os não nucleares se dividem em extratemáticos e periféricos. Desse modo, o evento, a frequência e as iterações são elementos do *frame* extratemáticos; e a duração, o modo, o lugar, a gravidade, o tempo e a entidade danificada são elementos do *frame* periféricos. Logo, devido à metodologia do projeto FrameNet, foi adotada essa mesma postulação para a realização da etapa de anotação das ocorrências. Veja abaixo os elementos do *frame* pertencentes ao *frame* de experiência de dano corporal:

¹ O projeto baseia-se nas definições dos *frames* em língua inglesa, pois visa construir para o português do Brasil a contraparte da rede semântica que, há dez anos, já vem sendo realizada para o inglês pelo projeto FrameNet, em Berkeley, na Califórnia, liderado pelo professor Charles Fillmore. Entretanto, não houve nenhum impedimento de que, na medida em que os dados fossem analisados, observasse algum contraste quanto à definição do *frame* em inglês em relação ao português; logo, a definição do *frame*, bem como as suas classificações, foi passível de modificações.

3.1. Elementos do frame:

3.1.1. Nucleares:

Parte_do_Corpo [PartC]

Tipo semântico: Parte_do_Corpo

O local do corpo do **experienciador** onde a lesão corporal se localiza.

Eu **cortei** meu dedo no envelope.

Experienciador [Exp]

Tipo semântico: sentiente

O ser ou entidade que é lesionado.

O esquilo **esfolou** seu rabo.

Então **eu** **me** **cortei** em um envelope.

3.1.2. Não nucleares

Evento [Evento]

O evento em que a lesão corporal ocorre.

Eu **quebrei** meu braço **em um acidente** .

Duração [Dur]

O período de tempo em que o **Experienciador** repetidamente se machuca.

Paulo deu pontapés no sofá **durante vários meses** antes de se acostumar com sua nova localização.

Frequência [Freq]

O número de vezes que o **Experienciador** tem a experiência de dano ao corpo.

Sue **bateu** seu dedão **três vezes** hoje.

Entidade_danificada [Ent_Dan]

O **Experienciador** se machuca em uma entidade danificadora.

Ele me perfurou, mas ele **acertou** seu punho **em meu nariz**.

Iterações [Int]

O número de vezes que o dano ocorre ao **experenciador**.

Em **machuquei** meu dedo mínimo nesses malabares **pela segunda vez** na última hora.

Manner [Manr]

Tipo semântico: maneira

Maneira na qual a lesão se encontra.

Lugar [Lugar]

Tipo semântico: relação_locativa

O lugar onde o evento danoso ocorre.

Ela **machucou** sua cabeça **no playground**

Gravidade [Grav]

A extensão na qual o **experenciador** é afetado pela lesão.

O suspeito **machucou** sua perna **gravemente** pulando a cerca.

Tempo

Tipo semântico: Tempo

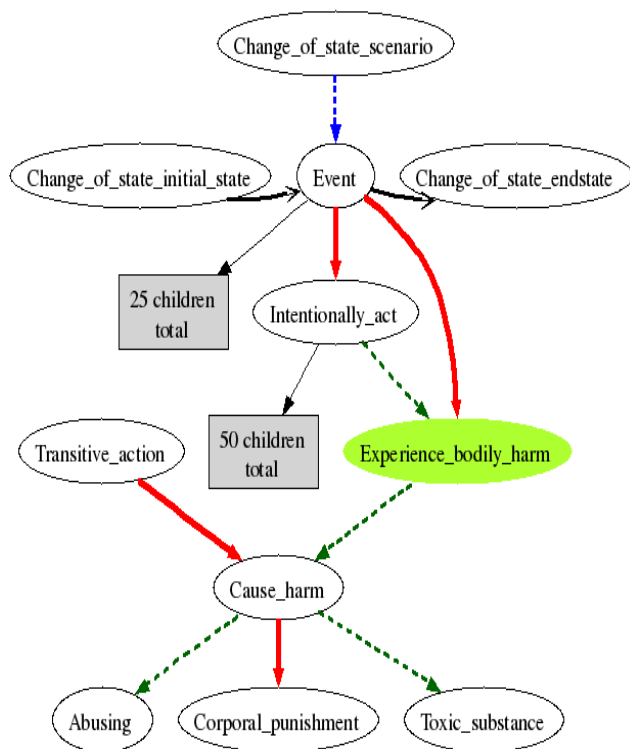
O tempo quando o evento danificador ocorre.

Então você **cortou** sua perna na moldura da janela **ontem**, certo?

Vale ressaltar que, no decorrer da etapa de rotulação dos elementos de *frames*, em algumas ocorrências, a unidade lexical “cortar” evocou o *frame cause harm*, traduzido como causar dano. Verificando, desse modo, que existe uma relação interframes, *frame* de experiência de dano corporal com *frame* causar dano. Tal relação, de acordo com a FrameNet, dá-se de forma que o primeiro é usado pelo segundo. Ou seja, o *frame* de experiência de dano corporal é subframe de causar dano. Observa-se, assim, que parte da cena evocada pelo *frame* usuário, nesse caso causar da-

no, é uma cena evocada pelo *frame* ancestral, o *frame* experiência de dano corporal¹.

Nesse sentido, constatou-se que há uma relação de herança entre o *frame* de experiência de dano corporal com o *frame* de evento. Sendo este o *frame* de herança e aquele o *frame* herdeiro, em outras palavras, o *frame* de experiência de dano corporal é um tipo de evento. Logo, tudo o que é verdadeiro no *frame* de evento também o é no *frame* de experiência de dano corporal. Por fim, ressalta-se que o *frame* de experiência de dano corporal usa o de ato intencional, *intentionally act*. A fim de compreender como que se processa essa relação interframes, segue o gráfico abaixo:



¹ Para a realização desta pesquisa, separaram-se as ocorrências que evocam o *frame* causar dano para, quiçá, futuras investigações e atentou-se apenas para as ocorrências do *frame* de experiência de dano corporal, já que essas são o foco do subprojeto.

Após a limpeza das ocorrências encontradas nos *corpora*, deu-se início ao processo de rotulação das ocorrências levantadas e, em seguida, o estabelecimento de padrões de valência dos predicadores lexicais. Segue abaixo os padrões de valência desenvolvidos na pesquisa:

3.2. Padrões de valência

3.2.1. 1º Padrão

Elementos do frame: Experienciador, Parte do corpo

Realizações sintáticas: Experienciador (Ext – SN), Parte do corpo (Obj –SN)

1) Bobbitter, a palavra já existente nos Estados Unidos, significa **CORTAR o pênis do amante**, inspirada naturalmente no affair Lorena Bobbit .

Camadas	CORTAR	o pênis	do amante
EF		Parte do corpo	experienciador
FG		Obj	Ext
TS		SN	SN
Verbo			

2) O braço. –Em Barcelonta? Ou contra os suecos? –Na casa dele, –com a mulher dele, e só por isso **me CORTOU o braço**. –Ah! –Depois disso ficamos... –amigos, fizemos as malas e... partimos os dois... –com o monge

3) Estava desequilibrado, –não foi a primeira vez que tentou se matar. –Porque matou Daniels. –Quase **lhe CORTOU a cabeça**, tirou-lhe o coração do peito! –Ninguém sonha com isso! Você viu tudo, não foi? – Você

4) Único aspecto a que eu tinha algum reparo a fazer era... –como é que podíamos fazer... –quando mandam **CORTAR a cabeça de João Batista**? –<i> – É algo que eu não tenho... </i> <i> – Delineado. </i> – <i> – Ainda

5) Baú cheio de tesouros. –A lenda diz que –ele deixou dois homens de sua tripulação vigiando. –Mas **CORTOU as pernas deles** pra que não pudessem fugir. – Isso é uma história estúpida. – Está bem. – Jim!

3.2.2. 2º padrão

Elementos do frame: Experienciador; Parte do Corpo

Realizações Sintáticas: Experienciador – INI-; Parte do Corpo – Obj – SN

6) Sabia que se acalmaria. –É só barbear amigo, fica calmo. – Adivinha. –Não se mova, não quero que **CORTE** a garganta. –Seu Filho da Puta ! Vamos! –Tira ele daí! Tira ele daí! –Que diabo... –Arthur! –Ei

Camadas	CORTE	a garganta	
EF		Parte do corpo	Experienciador INI
FG		Obj	
TS		SN	
Verbo			

7) Sabia? –Um homem quer uma garota que o escute. –Uma garota que cozinhe, costure e limpe. –Preferiria **CORTAR** o pescoço. **CORTAR** as veias. –Serei alguém importante. –Irei a Paris... –...vou estudar música e

8) Fisiológica, né? É assim uma cirurgia, não deixa de ser, e que muitas vezes aquela anestesia que se toma pode prejudicar. E você ter de **CORTAR** a barriga. E eu tenho a impressão que deve ter que **CORTAR** a musculatura também. Depois aquilo é tudo suturado, pode dar aderências

9) Não havia também, como se tratar, não havia jeito daquilo, ele ficou, ele tem uma mão aleijada, inutilizada, ficou sem, porque **CORTOU** nervo, tendão, veia, cortou tudo, por causa dos mosquitos e tudo isso. Nós "flitávamos" assim, botávamos aquele "flit", antes

3.2.3. 3º padrão

Elementos do frame: Experienciador (Conflação); Parte do Corpo

Realizações Sintáticas: Experienciador – INI-; Parte do Corpo – Obj – SN

10) Não Yoon-jin. Tae-yeon, faz favor. –Todo é culpa minha. – Farei tudo o que me peças. –Se desejar, **CORTAR** meus dedos. –Poderia arrancar-me os olhos. –Não, até poderia me tirar o coração. Faz favor. – Faz

Camadas	CORTE	meus dedos	
EF		Parte do corpo	Experienciador IN
FG		Obj	
TS		SN	
Verbo			

3.2.4. 4º padrão

Elementos do frame: Experienciador (conflação), Parte do corpo

Realizações sintáticas: Experienciador (Ext – SN), Parte do corpo (Obj –SN)

11) Especiais esteve muito mal. –E o Chuck e a Terry ficaram aleijados... –e... –Bem, os teus dedos foram CORTADOS –Aqui? – Heather, tu provavelmente estiveste meia acordada... –e vi os meus dedos serem cortados quando

Camadas	os teus dedos	foram	CORTADOS	
EF	Parte do corpo			Experienciador IN
FG	Obj			
TS	SN			
Verbo		Auxiliar		

12) Especiais esteve muito mal. –E o Chuck e a Terry ficaram aleijados... –e... –Bem, os teus dedos foram cortados. –Aqui? –Heather, tu provavelmente estiveste meio acordada... –e vi os meus dedos serem CORTADOS quando

3.2.5. 5º padrão

Elementos do frame: Experienciador (Conflação), Parte do corpo, Maneira

Realizações sintáticas: Experienciador (Ext – SN), Parte do corpo (Obj –SN), Maneira (Adv – SP)

13) Seu cérebro foi CORTADO em lâminas e usado em simulações por computador.

Camadas	Seu cérebro	foi	CORTADOS	em lâminas
EF	Parte do corpo			Maneira
FG	Obj			Adv
TS	SN			SP
Verbo		auxiliar		

3.2.6. 6º padrão

Elementos do frame: Experienciador (INI), Parte do corpo, Maneira

Realizações sintáticas: Experienciador INI , Parte do corpo (Obj – SN), Maneira (Adv – SP)

14) Só era possível ver a mesa operatória de longe, e nenhum close do pênis sendo CORTADO foi transmitido

Camadas	pênis	sendo	CORTADO	
EF	Parte do corpo			Experienciador INI
FG	Obj			
TS	SN			
Verbo		Auxiliar		

3.2.7. 7º padrão

Elementos do frame: Experienciador

Realizações sintáticas: Experienciador (Ext – SN)

15) Está me assustando. –Não, você está me assustando, Meryl. – O que você vai fazer? Me picar, fatar e CORTAR ? –Mil utilidades! – Faça alguma coisa! –O quê? O que disse? – Com quem está falando? – Nada. Eu não

Camadas	me	CORTAR
EF	experienciador	
FG	Ext	
TS	SN	
Verbo		

3.2.8. 8º padrão

Elementos do frame: Experienciador INI

Realizações sintáticas: Experienciador (INI)

16) E ele foi botar, mas com uma mão só quis pegar a janela que era de, de guilhotina, pra aguentar. Enterrou a mão pela janela, CORTOU tudo. Aí teve que ser operado, o médico me disse que não sabia nem se ele perderia a mão ou não, porque, começaram a fazer:

Camadas		CORTOU
EF	Experienciador	INI
FG		
TS		
Verbo		

3.2.9. 9º padrão

Elementos do frame: Experienciador (Ext– SN)

Realizações sintáticas: Experienciador (Ext-SN)

17) Aqueles pagãos selvagens sedentos de sangue! –Quantos morreram? –Foi uma família inteira, eles foram CORTADOS e mutilados. – Os pais e duas crianças. –A família Metz pegou a estrada Spearfish voltando para Minnesota

Camadas	eles	foram	CORTADOS
EF	Experienciador		
FG	Ext		
TS	SN		
Verbo		Auxiliar	

3.2.10. 10º padrão

Elementos do frame: Experienciador (Ext-SN)), Parte do corpo

Realizações sintáticas: Experienciador (Ext – SN), Parte do corpo (Obj –SN)

18) Você não atira em um delator. –Você faz ele sangrar, faz ele sofrer. –A língua de Clawson foi CORTADA fora. –Ele foi castrado e seu pênis foi enfiado em sua boca. –Isso é exatamente o que eu disse que

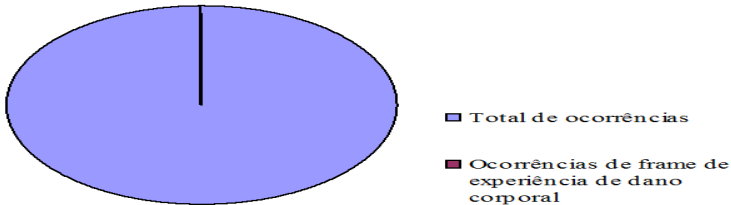
Camadas	A língua	Clawson	foi	CORTADA
EF	Parte do corpo	Experienciador		
FG	Obj	Ext		
TS	SN	SN		
Verbo			Auxiliar	

As tabelas abaixo mostram a distribuição do número de ocorrências levantadas com a unidade lexical “cortar” nos corpora correspondentes aos acervos eletrônicos AC/DC e Sketch Engine.

AC/DC	
Corpus	Total de ocorrências
ANCIB	10
ECI-EBR	92
NILC/São Carlos	2.258

SKETCH ENGINE	
Corpus	Total de ocorrências
Legendas de filmes	12.117
NURC-RJ	327

O gráfico indica que, do total de 14.804 ocorrências levantadas com a unidade lexical “cortar”, apenas em 18 encontrou-se o frame de experiência de dano corporal:



4. Discussão e conclusões

A partir da delimitação da noção de frame, enquanto categoria teórica, proposta nos estudos de Fillmore da Semântica Lexical, mesmo que ainda no modelo chomskyano, pretendia-se prever a sintaxe sentencial por meio da valência dos predicadores lexicais, tendência essa hoje muito explorada pelos linguistas nos estudos gramaticais. Mas foi, nos anos 80, com a afirmação da linguística cognitiva, que se compreendeu *frame* como um processo de conceptualização e categorização, ou seja, um tecido conceptual, cultural, cognitivo subjacente.

Visto que, conforme prevê essa corrente linguística, a linguagem é um processo cognitivo, é imprescindível, portanto, admitir que há conjuntos de conhecimentos sobre os quais a linguagem se realiza. Esses conjuntos são constituídos de estruturas mentais que podem ser estáveis ou transitórias. Nesse sentido, “[...] a linguagem funciona como operadora da conceptualização socialmente localizada através de um sujeito cognitivo, em situação comunicativa real, que produz significados como construções mentais, a serem sancionadas no fluxo interativo.” (SALOMÃO, 1999, p. 64). A linguagem, assim, funciona como um instrumento

cognitivo, decorrente da percepção visual e/ou do raciocínio, que aciona um conjunto de princípios operativos, cujas bases se fundamentam tanto na memória quanto no contexto da situação comunicativa (SALOMÃO, 1999).

Desse modo, a partir disso, pode-se admitir *frame* como um processo mental pelo qual organizamos o pensamento e categorizamos as coisas do mundo. Ou, em outras palavras, pode-se dizer que “as palavras pertencentes a um *frame* são representações lexicais de uma esquematização coerente da experiência o do conhecimento” (FILLMORE apud GAWRON, 2008).

É com base nessa noção de *frame* que esta pesquisa busca contribuir com a criação, a partir de evidências extraídas de *corpora* eletrônicos, de uma rede semântica aplicada ao léxico do português do Brasil, segundo os parâmetros do Projeto FrameNet Brasil, para que futuramente haja possibilidades de aplicação tecnológica. Como, por exemplo, a criação de um dicionário eletrônico, cujas entradas lexicais sejam definidas pelos *frames* a que correspondem.

Além disso, destaca-se, por último, que o projeto desperta grande interesse na área de Processamento da Linguagem Natural (PNL), já que os subcorpora constituídos no processo da descrição lexicográfica permitirão o treinamento de computadores para operar tarefas de tradução automática, extração de informação, paráfrases, sumariamento e tradução de textos.

REFERÊNCIAS BIBLIOGRÁFICAS

GAWRON, J.M. *Frame Semantics*. Manuscript, San Diego State University, 2008.

RUPPENHOFER, J. et al. *FrameNet II: Extended Theory and Practice*. 2006. Disponível em:

http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126. Acesso em: 07-2010.

SALOMÃO, Maria M. M. *FrameNet Brasil: Um trabalho em progresso*. In Calidoscópio, Vol. 7.2. 2009. Disponível em:

<http://www.franenetbr.ufjf.br>. Acesso em: 07-2010.

_____. A questão da construção de sentido e a revisão da agenda dos estudos da linguagem. In: *Veredas – Revista de Estudos Linguísticos*, v.3, n.1. Juiz de Fora: EDUFJF, 1999, p.61-79.