

DOI: <http://dx.doi.org/10.1590/1982-4017-160302-0316D>

CONVERGÊNCIAS ENTRE SEMÂNTICA DE *FRAMES* E LEXICOGRAFIA

Rove Chishman*

Universidade do Vale do Rio dos Sinos
Escola da Indústria Criativa
São Leopoldo, RS, Brasil

Resumo: O presente ensaio objetiva refletir sobre as convergências entre Semântica de Frames (FILLMORE, 1982, 1985) e Lexicografia. Para tanto, preveem-se dois vieses distintos, mas complementares para abordar o tema. Um deles diz respeito à interface da Semântica de Frames com a Lexicografia Computacional. Essa primeira perspectiva remete obrigatoriamente a tratar da Semântica de Frames e de sua contraparte computacional, a base de dados *FrameNet*, assim como a considerar o efervescente contexto em que se situa o projeto de Fillmore. O outro viés, através do qual se abordará a interface, leva em conta a aproximação não apenas da Semântica de Frames, mas da Semântica Cognitiva em termos mais amplos, com a Lexicografia tradicional. Dar-se-á destaque à noção de *frame* semântico nesse novo cenário de investigação.

Palavras-chave: Semântica de frames. Léxico computacional. Lexicografia.

1 INTRODUÇÃO¹

O presente ensaio objetiva refletir sobre as convergências entre Semântica de Frames (FILLMORE, 1982, 1985) e Lexicografia. Para tanto, prevêem-se dois vieses distintos, mas complementares para abordar o tema. Um deles diz respeito à interface da Semântica de Frames com a Lexicografia Computacional. Essa primeira perspectiva nos remete obrigatoriamente a tratar da Semântica de Frames e sua contraparte computacional, a base de dados *FrameNet*, assim como a considerar o efervescente contexto em que se situa o projeto de Fillmore. Salienta-se, nesse primeiro enfoque, o papel que os estudos semânticos e lexicais passaram a assumir no âmbito da Linguística Computacional e da Inteligência Artificial. O outro viés através do qual se abordará a interface leva em conta a aproximação não apenas da Semântica de Frames, mas da Semântica Cognitiva em termos mais amplos, com a Lexicografia tradicional. Trata-se de uma perspectiva ainda pouco explorada, mas que sinaliza para a pertinência em se aplicarem os preceitos da Semântica Cognitiva como elementos organizadores na estruturação dos dicionários. Dar-se-á destaque à noção de *frame* semântico nesse novo cenário de investigação.

* Professora do Programa de Pós-Graduação em Linguística Aplicada da Unisinos, Pesquisadora Produtividade CNPq (PQ 2). E-mail: rove@unisinos.br.

¹ Artigo desenvolvido como parte do projeto de pesquisa “Convergências entre Semântica de Frames e Lexicografia Computacional”, financiado com recursos do CNPq (Processo No. 311450/2015-6).

2 DISCUSSÃO

2.1 DOS DICIONÁRIOS AOS LÉXICOS COMPUTACIONAIS

Antes de nos atermos à apresentação da Semântica de *Frames* e sua contraparte computacional, a base de dados *FrameNet*, convém que se chame a atenção para um período em que os estudos semânticos e lexicais passaram a ser centrais no âmbito da Computação.

É na década de 1990 que questões de natureza semântica e lexical passaram a despertar o interesse dos estudiosos das áreas da engenharia da linguagem e da representação do conhecimento, assumindo papel central na construção de uma série de produtos para processamento da linguagem natural, tais como tradutores automáticos e programas de indexação e recuperação de informação.

Recebem importância, nesse período, os chamados *dicionários legíveis por máquina* (*machine-readable dictionaries* – MRDs), dicionários em formato eletrônico com possibilidade de serem processados e manipulados pelo computador. MRDs, em sua fase inicial, nada mais eram do que dicionários convencionais que se tornavam dicionários inteligentes, haja vista sua capacidade de processamento.

Litwovski (2003, p. 753-761) traz uma detalhada retrospectiva do surgimento e evolução desses recursos, destacando o *Longman's Dictionary of Contemporary English* (LDOCE), como um dos primeiros dicionários convencionais a serem usados como MRDs. Com os avanços tecnológicos, tais dicionários deixaram de ser transcrições fiéis dos dicionários impressos e passaram a contar com versões em formatos de mais fácil acesso, como é o caso da linguagem de marcação XML (*eXtensible Markup Language*).

Acompanhando a própria evolução da tecnologia nesse período, surgem também os dicionários construídos especialmente para uso no processamento da linguagem (também conhecidos como *PLN dictionaries*). Embora ambos possam ser processáveis pela máquina, os dicionários para uso exclusivo em PLN já são construídos tendo em vista exclusivamente o uso computacional.

Nessa mesma linha podemos compreender a denominação *dicionário tratável por máquina* (no inglês *machine tractable dictionary* ou MTDs) proposta por Wilks et al. (1996) para se referir ao resultado da transformação de um MRD com vistas a ser manipulado de forma mais apropriada tanto pela Linguística Computacional como pela Inteligência Artificial.

Esta sucinta contextualização nos permite compreender outras questões. Uma delas é a própria concepção alargada que os termos *dicionário* e *léxico* passaram a assumir. Segundo Litwovski (2003, p. 753-761), léxicos computacionais e dicionários incluem não apenas versões manipuláveis pelo computador de dicionários convencionais e thesaurus, como quaisquer compilações eletrônicas de palavras, sintagmas e conceitos, tais como listas de palavras, taxonomias, bases de dados terminológicas e ontologias. Destaca-se também que não se incluem nessa categorização ampla as versões eletrônicas dos dicionários, construídas apenas com o intuito de simples consulta na tela do computador por um usuário humano.

Outra questão relevante para a presente discussão diz respeito ao próprio status que as teorias semânticas passaram a adquirir, especialmente pela preocupação em atender às demandas provindas das aplicações computacionais. A Semântica Lexical representa, nessa visão revigorada, uma área de convergência entre semântica lexical, lexicografia computacional e semântica computacional.

Niremburg e Raskin (2004: p. 25) se referem às abordagens já delineadas tendo a aplicação computacional em mente como representando a visão de *demand-side lexical semantics* ou *ontology-driven theories*. Em face ao compromisso com a aplicação, trata-se de concepções teóricas que enfatizam as questões essenciais de significado, de maneira a não o dissociar do conhecimento de mundo. A profundidade e a cobertura da descrição do léxico dependem da arquitetura do sistema em desenvolvimento. O importante é construir modelos (isto é, ontologias) e pô-los em correspondência com o léxico, que passaria a conter conhecimento que não é exigido para a decodificação da realização sintática dos argumentos.

Essa visão de orientação ontológica contribuiria com a força expressiva adicional ao fornecer informação residual à estrutura predicado-argumento, como, por exemplo, especificações quanto às restrições de seleção e ao próprio conteúdo das unidades.

Contrastando com essa visão que mescla teoria e prática estão as abordagens que Niremburg e Raskin denominam de *supply-side lexical semantics* ou *syntax-driven lexical semantics*. Denominações como semântica gramatical ou semântica linguística também são empregadas como referindo-se a esta visão. Tais abordagens se atêm às noções de predicados e argumentos e a apresentar informações sobre os diferentes papéis temáticos que podem ser incluídas nas entradas lexicais. Entre as peculiaridades da visão sintaticamente orientada, destaca-se o interesse em formular teorias em semântica lexical que põem o foco de atenção na elegância formal, na economia de recursos descritivos e na ausência de exceções. Em consequência desse tipo de compromisso, fenômenos que não se submetem ao rigor almejado não são tratados. É o que normalmente ocorre com as questões de conteúdo lexical, que raramente são discutidas por teorias que seguem essa tradição. Estuda-se o que é possível.

Como exemplo dessa concepção, temos as investigações realizadas no âmbito do *MIT Lexicon Project*, cujos principais objetivos do projeto são o estudo de regularidades no comportamento de itens lexicais, em especial de verbos, e a descoberta de princípios que determinam o mapeamento entre semântica lexical e morfossintaxe (LEVIN, 2011).

2.2 A SEMÂNTICA DE FRAMES

A Semântica de *Frames* é uma das abordagens mais importantes do movimento nos estudos linguísticos conhecidos como Linguística Cognitiva. Tendo surgido no cenário linguístico nos fins da década de 70 e princípios da década de 80, a Linguística Cognitiva se estabelece como um paradigma que se desvincula de aspectos-chave presentes em abordagens formais em vigor na época, haja vista rejeitar a concepção de autonomia da linguagem, assim como o tratamento periférico dispensado a questões relacionadas ao sentido e ao uso.

Desse período inicial da Linguística Cognitiva, os linguistas George Lakoff, Ronald Langacker, Leonard Talmy, Gilles Fauconnier e Charles Fillmore figuraram como os principais expoentes que se empenharam em firmar este novo modelo para os estudos da linguagem, capaz de dar conta daquilo que era antes posto de lado.

São considerados centrais para a Linguística Cognitiva os seguintes aspectos: a) a centralidade do estudo das estruturas conceptuais; b) o motivacionismo semântico da gramática; c) a diluição de dicotomias como semântica/pragmática, significado linguístico/significado extralinguístico e sentido literal/sentido figurado; d) o compromisso com o experiencialismo; e e) a visão enciclopédica do significado.

No que tange à relação entre cognição humana e linguagem, a Linguística Cognitiva se afasta da perspectiva de modularidade da mente e defende o pressuposto de que a linguagem está intrinsecamente relacionada com os demais módulos cognitivos, assumindo a existência de aspectos cognitivos compartilhados pela linguagem e outras capacidades. Segundo Fauconnier e Turner (2003), capacidades cognitivas que possuem papel central na linguagem não são específicas a ela. Dentre essas capacidades, podemos incluir analogia, recursão, ponto de vista e perspectiva, alinhamento figura e fundo e integração conceptual. O fato de a Linguística Cognitiva promover o encontro de todos os níveis linguísticos, não mais negligenciando a semântica e a pragmática, possibilita compreendermos que o conhecimento é *enciclopédico*. Por exemplo, o significado de dicionário de “cachorro” se limita a incluir as propriedades de um cachorro que nos possibilitam distingui-lo de outros seres, enquanto o conhecimento enciclopédico carrega uma série de outras informações, como o fato de cachorros serem conhecidos, em determinados contextos socioculturais, como os melhores amigos do homem (CRUSE, 2006). Sendo assim, a Linguística Cognitiva relaciona a linguagem (e a forma como nos expressamos através dela) com as nossas experiências. É nesse contexto que surge a Semântica de *Frames* de Charles Fillmore.

A teoria, de acordo com Evans e Green (2006, p. 222, tradução nossa), “tenta descobrir as propriedades do inventário estruturado da linguagem associado com as palavras e considera quais consequências as propriedades desse sistema de conhecimento podem ter para um modelo semântico”². Um *frame* pode ser descrito como uma esquematização da experiência depositada em nossa memória de longo-prazo (EVANS; GREEN, 2006), elencando os elementos participantes de uma determinada experiência. De acordo com Fillmore (1982, p. 11), um *frame* é “qualquer sistema de conceitos relacionados de tal maneira que para entender qualquer um deles é preciso entender a estrutura que os comporta como um todo [...]”.

Um outro aspecto central para compreendermos a Semântica de *Frames* e os desdobramentos que interessam à nossa reflexão neste trabalho diz respeito ao compromisso da teoria com o empirismo. Empirismo, no contexto da Linguística Cognitiva, relaciona-se diretamente à visão experiencialista, a qual, num sentido lato, envolve aspectos culturais, históricos e sociais.

² No original: “[...] attempts to uncover the properties of the structured inventory of knowledge associated with words, and to consider what consequences the properties of this knowledge system might have for a model of semantics.”

Fillmore (1982) caracteriza a Semântica de *Frames* como empírica. Nas palavras do linguista, “A Semântica de *Frames* oferece um modo particular de se olhar para o significado das palavras [...]” (FILLMORE, 1982, p.11). Tal modo é caracterizado pela observação e análise das nossas experiências, resultantes da nossa interação com o ambiente. As palavras, para Fillmore, são *categorizações de experiência*, sustentadas por uma situação motivacional que ocorre em um plano de fundo de conhecimento decorrente da forma como nos posicionamos e percebemos o mundo ao nosso redor. Desse modo, os *frames* podem caracterizar um modo para se entender a razão pela qual determinada comunidade cria certas categorias de palavras, buscando explicar o significado de cada unidade lexical através do esclarecimento de tal motivação. Os *frames*, nesse sentido, estão intrinsecamente relacionados à cultura.

Tal relação existente entre o *frame* semântico de Fillmore e a cultura de cada um pode ser representada a partir de um *frame* de café da manhã (FILLMORE, 1982). Seguindo o linguista:

Entender essa palavra é entender a prática em nossa cultura de ter três refeições diárias, em horários mais ou menos convencionalmente estabelecidos do dia, e sendo que uma dessas refeições é aquela a ser feita no início do dia, após um período de sono, consistindo de um menu um tanto quanto único (cujos detalhes podem variar de comunidade para comunidade) (FILLMORE, 1982, p.118, tradução nossa).³

Seguindo o raciocínio de Fillmore, podemos dizer que compreender “café da manhã” significa compreender o papel dessa refeição na nossa sociedade. Podemos ir mais fundo: um café da manhã tradicionalmente americano contém ovos mexidos, bacon, panquecas. É possível que a ausência de algum desses elementos cause estranhamento em uma pessoa acostumada a essa realidade. No entanto, no Brasil, é igualmente possível que a presença dos mesmos elementos cause tal estranhamento. É importante ressaltar que o *frame* é o mesmo, uma vez que ele está ainda retratando a primeira refeição do dia, feita no início da manhã, após um período de sono, dentre outros aspectos que o caracterizam.

Ainda explorando o caráter empírico da Semântica de *Frames*, vale destacar o compromisso que Fillmore passa a assumir com as investigações baseadas em *corpora* para descrever a linguagem.

Dessa forma, há claramente uma refutação de abordagens intuitivas que desconsiderem o real uso linguístico dos falantes. Esses preceitos são retomados e aperfeiçoados na década de 1980, quando o autor consolida a Semântica de *Frames* como um programa de análise do significado a partir de situações reais de comunicação (FILLMORE, 1982; 1985).

Outro ponto de especial interesse para este trabalho diz respeito aos aspectos sintagmáticos do significado lexical que Fillmore passa a considerar. Cabe lembrar que foi no clássico *The case for cases* (1968) que Fillmore propôs, ainda no âmbito do

³ No original: “To understand this word is to understand the practice in our culture of having three meals a day, at more or less conventionally established times of the day, and for one of these meals to be the one which is eaten early in the day, after a period of sleep, and for it to consist of a somewhat unique menu (the details of which can vary from community to community).”

gerativismo, a Gramática de Casos. Tomando como base a teoria padrão, a ideia consistia em se dispor de uma representação mais profunda do que a já prevista pela teoria, com uma base sintática associada a elementos semânticos.

Nos anos de 1990, com a publicação de *Toward a Frame-Based Lexicon: The Semantics of RISK and its Neighbors* em co-autoria com Atkins, Fillmore (1992) traz uma análise descritiva minuciosa no que diz respeito às descrições valenciais que estruturam os *frames*. Dentre os principais resultados deste estudo está a evidência de que um dicionário em que se separam os sentidos com base em diferenças de padrões gramaticais não permite que o usuário verifique processos metafóricos e metonímicos relacionados a determinadas unidades lexicais (FILLMORE; ATKINS, 1992), daí a necessidade de se criar um recurso organizado em torno de *frames*.

É neste contexto que surge o projeto FrameNet.

2.3 A FRAMENET

A Plataforma *FrameNet*⁴ vem sendo desenvolvida no *International Computer Science Institute* (ICSI) em Berkeley (Califórnia), tendo entrado em operação em 1997 como uma contraparte aplicada da SF.

Trata-se de um recurso projetado para ser legível tanto por máquinas como por humanos, sendo, portanto, um MRD, conforme definição apresentada anteriormente. Nesse sentido, assemelha-se a outros projetos de bases de dados computacionais, tais como a *WordNet* (MILLER; FELLBAUM, 2007) e o *PropBank* (PALMER et al., 2005), apenas para citar os mais conhecidos, haja vista o compromisso com as aplicações em PLN.

Em artigo em que Salomão et al. (2013) refletem sobre a aplicação da Semântica de *Frames* na base da FrameNet, são apresentados o que os autores consideram como seus principais objetivos: a) a descrição das Unidades Lexicais (ULs), dos *frames* evocados e os Elementos de Frame (EFs); b) o uso de corpus para validar as descrições; c) a determinação das possibilidades combinatórias das ULs; d) a disponibilização de sentenças anotadas; e) a sumarização de padrões valenciais; e f) a representação das relações entre *frames*.

Desse modo, conforme sintetizam Atkins, Fillmore e Johnson (2003), a *FrameNet* é um recurso lexicográfico computacional que descreve propriedades semânticas e sintáticas de palavras em língua inglesa, a partir de *corpora* eletrônicos, apresentando essa informação através de uma plataforma *on-line*. Assim, o desenvolvimento do projeto embasa-se na análise de um vasto conjunto de textos autênticos de língua inglesa, do qual são extraídas as sentenças que ilustram as unidades lexicais.

Fillmore e Baker explicam que o método de pesquisa adotado pelo projeto

⁴ <<http://framenet.icsi.berkeley.edu>>.

[...] consiste em encontrar grupos de palavras cujas estruturas de *frames* podem ser conjuntamente descritas, devido ao fato de partilharem padrões e contextos esquemáticos comuns de expressões que podem se combinar com elas para formar frases ou sentenças maiores. Tipicamente, as palavras que partilham de um mesmo *frame* podem ser usadas como paráfrases umas das outras. As propostas gerais do projeto buscam oferecer descrições confiáveis de propriedades combinatórias sintáticas e semânticas de cada palavra do léxico e reunir informações sobre modos alternativos de se expressar conceitos dentro de um mesmo domínio conceptual. (FILLMORE; BAKER, 2010, p. 321, grifo nosso).⁵

Assim, o trabalho da *FrameNet* inclui o desenvolvimento do *frame* semântico, a extração de *corpus* e a anotação de sentenças exemplo. O desenvolvimento do *frame* tem como etapas a caracterização informal do tipo de entidade ou situação representada pelo *frame*, seleção dos elementos de *frame* e a construção de listas de palavras que pertencem ao *frame*. A extração de *corpus* tem por objetivo a verificação dos contextos sintáticos e semânticos dos constituintes das sentenças, anotando-se os exemplos de acordo com os EFs estabelecidos (FILLMORE; JOHNSON; PETRUCK, 2003).

Além da *Framenet* de Berkeley, estão em construção recursos lexicais baseados em *frames* para o hebraico (PETRUCK, 2009), o português brasileiro (SALOMÃO, 2009), o italiano (VENTURI et al., 2009) e o francês (PITEL, 2009). Os *frames* e elementos de *frame* da *FrameNet* também têm sido utilizados para anotação de *corpora* em alemão (BURCHARDT et al., 2009). Todos os recursos lexicais até aqui citados estão voltados à descrição de língua geral. No entanto, a *FrameNet* também serviu de inspiração para o desenvolvimento de recursos léxico-computacionais para linguagens especializadas. Um exemplo de aplicação dos *frames* para linguagens especializadas é o *Kicktionary* (SCHMIDT, 2009), base de dados lexicais multilíngue da linguagem do futebol e fonte de inspiração para o projeto *Kicktionary-Br* (CHISHMAN, 2009).

Importante destacar que estas *framenets* distintas, ainda que possam ser chamadas de léxicos computacionais, não foram desenvolvidas a partir dos preceitos da Lexicografia. O seu principal propósito é armazenar informações léxico-semânticas com base em *frames* em um formato que possa ser processado pelo computador. Trata-se, portanto, de bases de dados lexicais, e não dicionários propriamente ditos.

2.4 A SEMÂNTICA DE FRAMES E A LEXICOGRAFIA

Antes de tratar das possíveis convergências entre Linguística Cognitiva, em especial a teoria de nosso interesse – a Semântica de *Frames*, e a Lexicografia, é importante partir das possíveis convergências entre Lexicografia e Linguística em um sentido mais amplo.

⁵ No original: “The method of inquiry is to find groups of words whose frame structures can be described together, by virtue of their sharing common schematic backgrounds and patterns of expressions that can combine with them to form larger phrases or sentences. In the typical case, words that share a frame can be used in paraphrases of each other. The general purposes of the project are both to provide reliable descriptions of the syntactic and semantic combinatorial properties of each word in the lexicon, and to assemble information about alternative ways of expressing concepts in the same conceptual domain.”

Ainda que tal vinculação pareça natural e previsível, haja vista que os dicionários lidam com informação linguística, há quem chame a atenção para as dificuldades em estabelecer tais elos. O caráter fortemente aplicado da Lexicografia vem corroborar tal posição, que enfatiza que dicionários são artefatos (o que os distanciaria de um tratamento científico) e que ter embasamento em teoria linguística não é pré-requisito para uma prática bem-sucedida.

Atkins e Rundell (2008, p. 130) chamam a atenção para a situação inversa: teorias linguísticas têm aplicação direta no trabalho de planejamento e desenvolvimento de dicionários. Nesse cenário, a semântica lexical e outras abordagens filiadas ao estruturalismo estão entre as perspectivas teóricas que têm impactado a prática lexicográfica. Relações semânticas entre os itens lexicais, como hiponímia, sinonímia, meronímia e tipos variados de antonímia, e o tratamento da polissemia estão entre os fenômenos que têm sido abordados nos dicionários a partir do viés das teorias semânticas. Lyons (1969, 1977, apud ATKINS; RUNDELL, 2008) e Cruse (1986) estão entre os linguistas que têm inspirado o trabalho lexicográfico. Os princípios teóricos da chamada Lexicologia Combinatória Explanatória de Mel'cuk também têm sido aplicados na constituição de entradas lexicais. O *DiCoInfo – Dictionnaire fondamental de l'informatique et de l'Internet* - (LEROYER, 2007, 2011; L'HOMME et al., 2014) é um exemplo de tal aplicação ao fornecer informações sobre estrutura de actantes e circunstâncias nas entradas dos verbetes.

Geeraerts (2015), seguindo nessa mesma linha de pensamento, ressalta que a influência da perspectiva estruturalista já estava presente na obra de Zgusta (1971, apud Geeraerts, 2015), considerada como uma das mais influentes desse período. Dentre as perspectivas teóricas de orientação estruturalista que tiveram reflexo na prática lexicográfica, o autor chama a atenção para a teoria dos campos semânticos e sua aplicação em dicionários onomasiológicos⁶.

O autor também faz menção à semântica relacional e à semântica distribucional como abordagens estruturalistas que estão presentes em projetos lexicográficos. No caso da semântica relacional, é feita referência à aplicação que o projeto *WordNet* faz do conceito de relações de sentido. Na base de dados *WordNet*, nomes, verbos, adjetivos e advérbios são agrupados em conjuntos de sinônimos (comumente conhecidos como *synsets*). Como exemplo de projeto lexicográfico inspirado na semântica distribucional, Geeraerts faz menção ao trabalho de John Sinclair (ver Sinclair (1991)) com o desenvolvimento do *Collins Cobuild English Language Dictionary* (1987), para o qual foi compilado um corpus de 20 milhões de palavras. O conceito central desta perspectiva repousa sobre a noção de colocação, definida como uma relação lexical entre duas ou mais palavras que têm a tendência de aparecerem juntas em texto corrido.

No que diz respeito à aplicação da Linguística Cognitiva ao estudo dos dicionários, pode-se afirmar que as iniciativas são ainda tímidas. À primeira vista, Lexicografia e Linguística Cognitiva parecem não ter muitos pontos em comum ao ponto de favorecer

⁶ Dicionários que organizam o vocabulário não em ordem alfabética, mas com base em temáticas específicas e na associação semântica entre as palavras.

uma projeção de uma sobre a outra. Enquanto a Lexicografia tem seguido princípios estabelecidos por décadas e até por séculos, a Linguística Cognitiva, comparativamente, é um campo da Linguística que emergiu há poucas décadas. Geeraerts (1997, 2001, 2006, 2007) é um dos poucos teóricos da perspectiva cognitivista que tem discutido aspectos referentes à Lexicografia.

Referindo-se ao impacto da semântica cognitiva na lexicografia, Geeraerts (2015) dá destaque a três perspectivas: a semântica de *frames*, o estudo da metáfora conceptual e metonímia e a teoria dos protótipos. No que tange à semântica de *frames*, o autor chama a atenção para o projeto *FrameNet* e para o uso de corpus como a principal fonte de evidência empírica para a análise baseada em *frames*.

Os estudos sobre metáfora conceptual e metonímia, por sua vez, sinalizam para formas de lidar com as associações entre os sentidos de itens lexicais que vão além da prática comum nos dicionários. Como exemplo de iniciativas envolvendo as metáforas conceptuais, Geeraerts cita o *Macmillan English Dictionary for Advanced Learners* (2007), que incorpora “caixas de metáfora”, mostrando metáforas conceptuais subjacentes às expressões.

A teoria dos protótipos e a própria visão de polissemia e categorização que a Linguística Cognitiva defende impõem um desafio para uma perspectiva estruturalista de estudar o léxico. Em uma perspectiva que põe o foco sobre os sentidos como parte de um sistema linguístico mais ou menos autônomo, muitos dos fenômenos que são destacados pela semântica cognitiva ficam descartados.

Outro estudo que explora a aplicação dos preceitos da Linguística Cognitiva à prática lexicográfica é o de Carolin Ostermann (2015). A autora propõe uma nova abordagem para a lexicografia fazendo uso, em especial, da semântica cognitiva. À primeira vista, defende Ostermann, lexicografia e linguística cognitiva não parecem ter pontos em comum que favoreceriam a projeção de uma sobre a outra; contudo, o compromisso que a Linguística Cognitiva tem em descrever a linguagem de acordo com o modo como as pessoas percebem e conceptualizam o mundo, com a linguagem não sendo vista como uma faculdade isolada, mas relacionada ao mundo, parece adequada para a composição das estruturas dos dicionários. Segundo a autora, a Linguística Cognitiva pode auxiliar a lidar com problemas clássicos, como linearização, multiplicidade de sentidos, definições, metáforas e metonímias, além de partir de uma concepção psicologicamente plausível.

No que se refere especificamente ao emprego da noção de *frame* na prática lexicográfica, interesse central deste trabalho, Martin (1997, apud OSTERMANN, 2015) enfatiza a força descritiva e explanatória dos *frames*, haja vista que organizam e apresentam o conhecimento de mundo. Dentre as formas como os *frames* poderiam ser adicionados às entradas de um dicionário, o autor faz menção a uma possível organização temática ou até a eliminação dos sentidos do dicionário.

A proposta apresentada por Ostermann considera as sentenças de exemplos que fazem parte de uma entrada de dicionário como o *locus* para a aplicação da noção de

frame. Independente de sua função de mostrar a cabeça⁷ de verbete em um contexto específico e oferecer colocações, as sentenças evocam uma situação e oferecem espaço suficiente para descrições elaboradas sem as restrições de um formato específico.

Langacker (2005, apud OSTERMANN, 2015) e Atkins (1995, apud OSTERMANN, 2015) também reconhecem o papel que as sentenças de exemplos ocupam. Segundo Atkins, *as sentenças constituem uma parte integral da descrição do significado e devem de forma sistemática exemplificar todos os fatos relevantes do frame*. As seções de exemplos, sob tal óptica, adicionariam uma representação de um frame às entradas de dicionário dos itens lexicais participantes.

O desenvolvimento do *Dicionário Field: expressões do futebol*⁸ (CHISHMAN et al., 2014) foi guiado, ainda que de forma embrionária, pelo propósito de aplicar a noção de frame na estrutura do dicionário. Diferentemente do projeto *FrameNet*, o Field procurou ajustar os preceitos da Semântica de *Frames* à prática lexicográfica tradicional, o que significa contemplar requisitos como, por exemplo, o perfil do usuário.

O *Dicionário Olímpico*⁹ (CHISHMAN, 2015), por sua vez, vem sendo desenvolvido tendo como mote uma proposta similar à de Ostermann; contudo, ao invés de explorar as sentenças de exemplos como o lócus para a representação do frame, o dicionário para os esportes olímpicos põe ênfase na própria glosa do esporte e no princípio fillmoriano de que saber o significado de uma palavra implica saber o frame evocado por tal palavra.

3 CONSIDERAÇÕES FINAIS

Propusemo-nos, neste ensaio, a refletir sobre as convergências entre a Semântica de *Frames* e a Lexicografia. Iniciamos nosso percurso chamando a atenção para os léxicos computacionais, em especial destacando a distinção entre dicionários legíveis (ou tratáveis) por máquina e dicionários legíveis na máquina.

Partindo deste esclarecimento sobre os léxicos computacionais, apresentamos a Semântica de *Frames* e sua vinculação à Linguística Cognitiva, destacando as características que vão justificar o emprego desse arcabouço teórico na prática lexicográfica.

Tendo situado a Semântica de *Frames*, em seguida tratamos de apresentar duas possibilidades de aplicação da teoria. A primeira traz o projeto *FrameNet*, enfatizando as características que o situam como léxico computacional e indicam a interface com a Linguística Computacional. A segunda explora a interface da Semântica de *Frames* com a Lexicografia tradicional, destacando a pertinência de explorar não apenas a noção de *frames*, mas outros conceitos da Linguística Cognitiva.

⁷ Cabeça de verbete equivale à entrada de verbete. Optou-se pela forma *cabeça de verbete* em consideração à expressão original em inglês *headword*, conforme Atkins e Rundell (2008).

⁸ <www.dicionariofield.com.br>.

⁹ <www.dicionariolimpico.com.br>.

Por fim, citamos dois projetos do grupo *SemanTec*¹⁰ que têm procurado aprofundar formas de aproximar a Semântica de *Frames* e a Lexicografia. Nesses projetos, nosso propósito foi apenas destacar uma das formas em que as duas áreas podem convergir, que é a noção de *frame* como princípio organizador de uma obra lexicográfica. Chamou-nos a atenção nesses empreendimentos a importância da noção de significado enciclopédico para a eficácia da compreensão dos sentidos.

Conforme já se destacou, não apenas a Semântica de *Frames*, mas também a semântica cognitiva em um sentido amplo, têm muito a oferecer para a prática lexicográfica. Categorização, polissemia, categorias radiais, metáforas conceptuais são tópicos que muito têm a agregar à prática dicionarística.

REFERÊNCIAS

- ATKINS, S.; FILLMORE, C.; JOHNSON, C. Lexicographic Relevance: Selecting Information from Corpus Evidence. In: *International Journal of Lexicography*, v. 16, n. 3, p. 251-280, 2003.
- ATKINS, S.; RUNDELL, M. *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press, 2008.
- BAKER, C. *FrameNet: A Knowledge Base for Natural Language Processing*. ACL 2014, v. 1929, p. 1-5, 2014.
- BURCHARDT, A.; ERK, K.; FRANK, A.; KOWALSKI, A.; PADÓ, S.; PINKAL, M. Using FrameNet for the semantic analysis of German: Annotation, Representation, and Automation. In: BOAS, H. C. (Ed.). *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Berlin/New York: Mouton de Gruyter, 2009. p. 209-244.
- CHISHMAN, R. L. O. *FrameCorp: uma proposta de anotação semântica aplicada à temática do futebol*. Projeto de Pesquisa apresentado ao CNPq, 2009.
- _____. *Convergências entre Semântica de Frames e Lexicografia Computacional*. Projeto de Pesquisa apresentado ao CNPq, 2015.
- CHISHMAN, R. L. O. et al. Field: Dicionário de expressões do futebol: um recurso lexicográfico baseado no aporte teórico-metodológico da semântica de frames e da linguística de corpus. *Signo*, v. 39, p. 25-35, 2014.
- CRUSE, A. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.
- _____. *A Glossary of Semantics and Pragmatics*. Edinburg: Edinburgh University Press, 2006.
- EVANS, V.; GREEN, M. *Cognitive Linguistics: An Introduction*. Edinburg: Edinburgh University Press, 2006.
- FAUCONNIER, G.; TURNER, M. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books, 2003.
- FILLMORE, C. J. The case for cases. In: BACH, E.; HARMS, R. (Eds.) *Universals in Linguistic Theory*. New York: Holt, Rinehart & Winston, Inc., 1968.

¹⁰ O grupo de pesquisa *SemanTec* (forma acrônima de Semântica e Tecnologia) é credenciado pelo CNPq desde 2000 e coordenado pela Profª. Dra. Rove Chishman. As investigações realizadas pelo grupo integram duas áreas do conhecimento: a Linguística, mais precisamente a Semântica Cognitiva, e a Informática, com ênfase nas investigações voltadas para a Lexicografia Computacional e o uso de *corpus* eletrônico (como roga a Linguística de *Corpus*) em situações de pesquisa. Mais informações em <www.projeto.unisinos.br/semantec>.

- _____. Frame Semantics. In: THE LINGUISTIC Society of Korea (Ed.). *Linguistics in the morning calm*. Seoul: Hanshin, 1982. p. 111-138.
- _____. Frames and the Semantics of Understanding. *Quaderni di Semantica*, v. 6, n.2, p. 222-254, 1985.
- _____; JOHNSON, C.; PETRUCK, M. Background to FrameNet. *International Journal of Lexicography*. Vol. 16.3, p. 235-250, 2003.
- _____. Frame Semantics. In: GEERAERTS, D. (Ed.) *Cognitive Linguistics: Basic readings*. Berlin: Mouton de Gruyter, 2006. p. 373-400.
- _____; ATKINS, B. T. S. Towards a frame-based lexicon: The semantics of RISK and its neighbors. In: LEHRER, A.; KITTAY, A. (Eds.) *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*. Hillsdale: Lawrence Erlbaum Associates, 1992. p. 75-102.
- _____; BAKER, C. A frames approach to semantic analysis. In: HEINE, B.; NARROG, H. (Eds.). *The Oxford Handbook of Linguistic Analysis*. New York: Oxford University Press, 2010. p. 313-339.
- GEERAERTS, D. *Diachronic Prototype Semantics*. A contribution to historical lexicology. Oxford: Clarendon Press, 1997.
- _____. On measuring lexical convergence. In: Soares da Silva, Augusto (Ed.), *Linguagem e Cognição. A Perspectiva da Linguística Cognitiva*. Braga: Associação Portuguesa de Linguística, 2001, p. 51-61.
- _____. *Cognitive Linguistics: Basic Readings*. Berlin: Mouton, 2006.
- _____. *Words and Other Wonders*. Papers on Lexical and Semantic Topics. Berlin/New York: Mouton de Gruyter, 2007.
- _____. Foreword. In: KOCKAERT, HENDRIK J.; STEURS, FRIEDA (Ed.), *Handbook of Terminology: Volume 1*. Amsterdam: John Benjamins Publishing Company, 2015, p. xiv-xvi.
- LEROYER, P. Bringing Corporate Dictionary Design into Accord with Corporate Image. In: GOTTLIEB, H.; MOGENSEN, J. (Eds.). *From Words to Messages and Back Again*, 2007, p. 109-117.
- _____. Change of Paradigm: From Linguistics to information science and from dictionaries to lexicographic information tools. In: FUERTES-OLIVERA, P., BERGENHOLTZ, H. (Eds.). *E-lexicography: The Internet Digital Initiatives and Lexicography*. London and New York: Continuum, 2001. p. 121-140.
- LEVIN, B. *The Lexicon Project and its legacy*, 2011. Disponível em: <<http://web.stanford.edu/~bclevin/levin11mitslides.pdf>>. Acesso em: 30 jun. 2016.
- LITKOWSKI, K. *Integrating Data from The Proposition Project into FrameNet*. CL Research, Damascus, MD, 2003.
- L'HOMME et al. Why lexical semantics is important for e-lexicography. *International Journal of Lexicography*, v. 27, n. 4, p. 1-8, Dec. 2014.
- MILLER, G.; FELLBAUM, C. WordNet Now and Then. *Language Resources and Evaluation*, v. 41, n. 2, p. 209-214, 2007.
- NIRENBERG, S.; RASKIN, V. *Ontological Semantics*. Cambridge, MA: MIT Press, 2004.
- OSTERMANN, C. *Cognitive Lexicography*. Boston: De Gruyter Mouton, 2015.
- PALMER, M.; GILDEA, D.; KINGSBURY, P. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, v. 31, n. 1, p. 71-105, March, 2005.
- PETRUCK, M. Typological considerations in constructing a Hebrew FrameNet. In: BOAS, H. C. (Ed.). *Multilingual FrameNets in computational lexicography: Methods and applications*. Berlin/New York: Mouton de Gruyter, 2009. p. 163-182.
- PITEL, G. Cross-lingual Labeling of Semantic Predicates and Roles: A Low-resource Method Based on Bilingual L(atent) S(ematic) A(nalysis). In: BOAS, H. C. (Ed.). *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Berlin/New York: Mouton de Gruyter, 2009. p. 245-284.
- SALOMÃO, M. FrameNet Brasil: um trabalho em progresso. *Calidoscópico (UNISINOS)*, v. 7, p. 171-182, 2009.
- SALOMÃO, M. et al. A linguística cognitiva encontra a linguística computacional: notícias do projeto Framenet Brasil. *Cadernos de Estudos Linguísticos*, Campinas, v. 55, n.1, p. 1-28, 2013.

SCHMIDT, T. The Kicktionary – A Multilingual Lexical Resource of Football Language. In: BOAS, Hans C. (Ed.): *Multilingual Framenets in Computational Lexicography*. New York: de Gruyter, 2009. p.101-134.

SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

VENTURI, G.; LENCI, A.; MONTEMAGNI, S.; VECCHI, E.; SAGRI, M.; TISCORNIA, D.; AGNOLONI, T. Towards a FrameNet Resource for the Legal Domain. In: *Proceedings of the Third Workshop on Legal Ontologies and Artificial Intelligence Techniques, Workshop on Semantic Processing of Legal Texts (LOAIT '09)*. Barcelona, Spain, 8 June, 2009.

WILKS, Y. et al. The Companions: Hybrid-World Approach. In: *IJCAI09 Workshop on knowledge and reasoning in practical dialogue systems*, 1996.

Recebido em: 04/07/16. Aprovado em: 22/07/16.

Title: *Similarities between Frame Semantics and Lexicography*

Author: *Rove Chishman*

Abstract: *This essay aims to reflect on the similarities between Frame Semantics (FILLMORE 1982, 1985) and Lexicography. To this proposal, two distinct biases but complementary are considered. One of them concerns the interface between Frame Semantics and Computational Lexicography. This first perspective obligatorily refers to the Frame Semantics and its computer counterpart, the database FrameNet, as well as to consider the effervescent context in which lies the Fillmore project. The other bias through which to address the interface takes into account the approach of not only Frame Semantics, but also the Cognitive Semantics in broader terms, with traditional lexicography. We will highlight the notion of semantic frame in this new research scenario.*

Keywords: *Frame Semantics. Computational lexicon. Lexicography.*

Título: *Convergencias entre Semántica de Frames y Lexicografía*

Autor: *Rove Chishman*

Resumen: *Este artículo tiene el objetivo de reflejar sobre convergencias entre Semántica de Frames (FILLMORE, 1982, 1985) y Lexicografía. Para ello, se prevén dos puntos de vista distintos, pero complementares para abordar el tema. Uno de ellos es con respecto a interface de la Semántica de Frames con la Lexicografía Computacional. Esa primera perspectiva remete obligatoriamente a tratar de la Semántica de Frames y de su contraparte computacional, la base de datos FrameNet, también a considerar el efervescente contexto en que se sitúa el proyecto de Fillmore. El otro punto de vista es a través del cual se aborda la interface y considera la aproximación no solo de la Semántica de Frames, pero de la Semántica Cognitiva en términos más amplios, con la Lexicografía tradicional. Se da destaque para la noción de frame semántico en ese nuevo escenario de investigación.*

Palabras-clave: *Semántica de Frames. Léxico computacional. Lexicografía.*



Este texto está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.