

Aplicando a Semântica de *Frames* na descrição do Direito brasileiro

Anderson Bertoldi (PPGLA – UNISINOS)

Rove Chishman (PPGLA – UNISINOS / CNPq)

RESUMO: Este trabalho tem como objetivo discutir questões de equivalência de tradução e correspondência conceptual de termos jurídicos em inglês e português. A estrutura conceptual evocada pelo termo jurídico em inglês é contrastada com a estrutura conceptual evocada por seu equivalente em português. Para o contraste da estrutura conceptual evocada pelo termo jurídico em inglês, são utilizados os *frames* semânticos da FrameNet. As diferenças conceptuais entre os *frames* jurídicos não estão relacionadas à língua, mas à cultura jurídica de cada país. Assim, o emprego dos *frames* da FrameNet para construção de recursos lexicais em diferentes línguas pode causar distorções conceptuais significativas. A discussão aqui apresentada tem origem no trabalho de desenvolvimento de um recurso léxico-computacional baseado em *frames* para a linguagem jurídica brasileira.

PALAVRAS-CHAVE: Semântica de Frames; Equivalência; Recursos lexicais multilíngues.

ABSTRACT: This paper presents a discussion on lexical equivalence and conceptual correspondence of legal terms in English and Portuguese. The conceptual structure evoked by legal terms in English is contrasted to the conceptual structure evoked by their equivalents in Portuguese. FrameNet's semantic frames are used to contrast conceptual structures of legal terms in English and Portuguese. This paper concludes that differences in conceptual structures of legal frames are related to culture, not to lexical equivalence; then, the development of multilingual lexical resources for legal language based on lexical equivalence may cause some conceptual mismatches. The discussion presented in this paper is motivated by the development of frame-based lexical resource for the Brazilian legal language.

KEYWORDS: Frame Semantics; Equivalence; Multilingual lexical resources.

Introdução

A Semântica de *Frames* (FILLMORE, 1975; 1977; 1982; 1985) já tem uma considerável tradição nos estudos multilíngues. Os conceitos de “cenas” e “*frames*”, apresentados por Fillmore (1977), têm sido utilizados nos estudos de tradução desde a década de 80 (VANNEREM; SNELL-HORNBY, 1986; SNELL-HORNBY, 1988; 2005; VERMEER; WITTE, 1990). Mais recentemente, a Semântica de *Frames* tem despertado o interesse de estudiosos dedicados ao desenvolvimento de recursos léxico-computacionais multilíngues, tais como *framenets* multilíngues e *corpora* com anotação de *frames* semânticos. A coletânea de trabalhos editada por Boas (2009) exemplifica bem o emprego da Semântica de *Frames* para a criação de recursos léxico-computacionais multilíngues baseados em *frames*.

Este trabalho tem como objetivo discutir questões de equivalência de tradução e correspondência conceptual envolvidas no desenvolvimento de um recurso léxico-computacional baseado em *frames* para a linguagem jurídica brasileira, com equivalentes em inglês. A metodologia de trabalho adotada, até o presente momento, para a construção desse recurso de linguagem especializada pode ser resumida em três etapas. Primeiramente, são identificados na FrameNet (FILLMORE; JOHNSON; PETRUCK, 2003) *frames* relacionados

ao domínio jurídico. Em segundo, procura-se no universo jurídico brasileiro um evento jurídico correspondente para o evento jurídico descrito pelo *frame* semântico da FrameNet. Em terceiro lugar, havendo no sistema jurídico brasileiro algum evento correspondente ao evento descrito pelo *frame* identificado na FrameNet, os eventos jurídicos são contrastados para identificar-se o nível de correspondência entre eles. Por fim, são criados *frames* para descrever os eventos jurídicos brasileiros e, em caso de correspondências com os eventos jurídicos identificados na FrameNet, são indicadas unidades lexicais em inglês equivalentes às unidades lexicais em português¹.

Nesse trabalho de verificação de correspondência conceptual e identificação de equivalentes em inglês para os termos jurídicos em português, percebeu-se que os *frames* jurídicos apresentam diversos níveis de correspondência. Essa correspondência pode se dar tanto em nível lexical (a equivalência das unidades lexicais em português e inglês), quanto no nível conceptual (o conhecimento especializado evocado pelas unidades lexicais em português e em inglês). A falta de correspondência conceptual entre *frames* jurídicos pode ser problemática para aplicações computacionais automáticas, como anotação automática de *corpus* com *frames* semânticos (PADÓ; LAPATA, 2005; PADÓ, 2007; TONELLI; PIANTA, 2008), uma vez que essas aplicações partem da identificação de equivalentes no nível lexical, o que nem sempre significa uma correspondência no nível conceptual.

Este trabalho discute como unidades lexicais, que em dicionários especializados bilíngues são dadas como equivalentes, evocam conhecimentos especializados distintos. Para tanto, este trabalho está estruturado em três seções. Na seção 1, discute-se a aplicação da Semântica de *Frames* no desenvolvimento de recursos lexicais multilíngues. Na seção 2, trata-se da aplicação da Semântica de *Frames* para o desenvolvimento de recursos lexicais multilíngues para o domínio jurídico. Na seção 3, aborda-se a correspondência de *frames* jurídicos de línguas e culturas jurídicas distintas. Na seção 4, discute-se como as divergências entre *frames* jurídicos podem afetar o desenvolvimento de recursos especializados para o domínio jurídico.

1. A Semântica de *Frames* e suas aplicações multilíngues

A Semântica de *Frames* nasce a partir de um conceito muito discutido na década de 70, o *frame* (MINSKY, 1974; GOFFMAN, 1974). Inicialmente, Fillmore (1975) faz uma distinção entre os conceitos de “cena” e “*frame*.” A cena não seria apenas uma cena visual, mas todo um conjunto de tipos familiares de transações interpessoais, cenários padrões definidos culturalmente, estruturas institucionais, crenças humanas, ações, experiências e imagens. Em contrapartida, o *frame* seria um sistema de escolhas linguísticas, sejam palavras, regras ou categorias gramaticais, associadas à determinada instância prototípica de uma cena.

A distinção entre cena como estrutura cognitiva e *frame* como estrutura linguística é posteriormente abandonada (FILLMORE, 1982; 1985). Segundo Fillmore (1982), “pelo termo ‘*frame*’ eu tenho em mente qualquer sistema de conceitos relacionados de tal forma que para entender qualquer um deles você tem que entender toda a estrutura na qual ele se encaixa

¹ Para este trabalho, utilizou-se o *corpus* NILC/São Carlos. Esse *corpus* contém textos em português brasileiro de diferentes gêneros. O *corpus* NILC/São Carlos é composto de textos jornalísticos, didáticos, epistolares e de redações de alunos (NUNES ET AL., 1996). Um *corpus* específico de linguagem jurídica está sendo desenvolvido no contexto do projeto Tecnologias Semânticas e Sistemas de Recuperação de Informação Jurídica, coordenado pela professora Dra. Rove Chishman e financiado pelo Edital N° 020/2010/CAPES/CNJ (CNJ Acadêmico). Um dos objetivos para o desenvolvimento de um *corpus* especializado em linguagem jurídica é a sua utilização no desenvolvimento de um recurso lexicográfico organizado por *frames* semânticos que descreva o significado das unidades lexicais do domínio jurídico brasileiro.

(...)” (p.111).² Para a Semântica de *Frames*, as palavras têm a capacidade de “evocar” todo um conhecimento de mundo que é organizado através de uma estrutura cognitiva chamada de *frame*: “Um *frame* é evocado pelo texto se alguma forma ou padrão linguístico é convencionalmente associado com o *frame* em questão” (FILLMORE, 1985, p.232).³

A Semântica de *Frames* é a base teórica que alicerça o desenvolvimento da FrameNet. A FrameNet foi concebida como um recurso lexical monolíngue, descrevendo o significado das unidades lexicais da língua inglesa com relação aos *frames* semânticos que cada unidade lexical evoca. No entanto, o desenvolvimento de projetos semelhantes, como a German FrameNet (BOAS, 2002; 2005), a Spanish FrameNet (SUBIRATS; PETRUCK, 2003) e a Japanese FrameNet (OHARA ET AL., 2003), estimulou a discussão sobre o uso da FrameNet como ponto de conexão entre bases de dados lexicais de línguas diferentes, criando um léxico multilíngue (BOAS, 2005).

Padó (2007) propôs o uso dos *frames* semânticos para o desenvolvimento automático de recursos léxico-computacionais multilíngues. Padó (2007) partiu do pressuposto de que os *frames* representariam um nível conceitual, sendo independentes das peculiaridades lexicais. Assim, os *frames* poderiam ser utilizados como ponto de conexão entre os itens lexicais, superando o problema da falta de paralelismo entre as línguas.

Além dos recursos lexicais já citados, 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). O Kicktionary é uma base de dados lexicais multilíngue da linguagem do futebol, em inglês, alemão e francês. Outra aplicação da FrameNet em domínios especializados tem sido a anotação de textos legais em italiano (VENTURI, 2011). Pode-se citar como outro exemplo de aplicação da Semântica de *Frames* para domínios especializados o trabalho de criação de um recurso léxico-computacional para a linguagem do Direito brasileiro (BERTOLDI; CHISHMAN, 2012).

2. A Semântica de *Frames* aplicada ao domínio jurídico

Poucos projetos têm se dedicado ao desenvolvimento de recursos léxico-computacionais multilíngues para o domínio jurídico. Talvez o projeto mais conhecido até então seja o LOIS⁴ (*Lexical Ontologies for Legal Information Sharing*), projeto de pesquisa realizado entre os anos de 2004 e 2005 e coordenado pelo *Istituto di Teoria e Tecnica della Informazione Giuridica* (ITTIG-CNR).⁵ O LOIS foi desenvolvido com o objetivo de facilitar a extração de informação em bases de dados jurídicos através da construção de uma *wordnet* jurídica multilíngue (DINI ET AL., 2005). Esse projeto multilíngue aplicado ao Direito foi

² “By the term ‘frame’ I have in mind any 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 (...)”. Todas as traduções dos originais em inglês citados neste trabalho são de responsabilidade dos autores.

³ “A frame is evoked by the text if some linguistic form or pattern is conventionally associated with the frame in question.”

⁴ www.loisproject.org

⁵ <http://www.ittig.cnr.it/Ricerca/Unita.php?Id=70&T=E>

inspirado em dois outros projetos de língua geral: a WordNet (MILLER, 1995; FELLBAUN, 1998) e a EuroWordNet (VOSSEN, 1998).

O modelo de agrupamento de sinônimos em *synsets*, apresentado pela WordNet, quando aplicado ao domínio jurídico, facilitaria a conexão de termos jurídicos, utilizados majoritariamente por especialistas, e expressões não especializadas, utilizadas majoritariamente pelo público leigo. Essa conexão de variantes sócio-dialetais facilitaria ao leigo o acesso à informação jurídica. Um exemplo dessa variação sócio-dialetal é o termo *locazione di immobile* (*locação de imóvel*) e seu equivalente não especializado *affito* (*aluguel*) (SAGRI; TISCORNIA, 2004). Já o modelo de integração de bases de dados multilíngues desenvolvido pelo projeto EurowordNet garantiria a interconexão das *wordnets* de diferentes línguas (CURTONI ET AL., 2005). Uma base de dados com essa estrutura permitiria o acesso, por leigos e especialistas, a leis e decisões judiciais dos vários países que integram a União Europeia.

A opção por utilizar a Semântica de *Frames* como paradigma teórico para o desenvolvimento de um recurso léxico-computacional especializado, e não um modelo de relações léxico-conceituais, como o paradigma teórico que norteia as *wordnets*, deve-se à possibilidade de mapeamento dos termos relacionados a um determinado evento jurídico que se está descrevendo. Conforme os princípios teóricos da Semântica de *Frames*, entender o significado de uma palavra requer o conhecimento de todos os conceitos que estão relacionados a essa palavra. Em outras palavras, entender o significado de *judgar* requer do falante o conhecimento de outros conceitos que estão relacionados ao verbo *judgar*, como *juiz*, *réu*, *delito* e *sentença*. Assim, a Semântica de *Frames* permite a relação dos participantes e dos eventos jurídicos por meio dos *frames* semânticos. Um recurso léxico-computacional baseado em *frames* pode ser útil para aplicações em recuperação de informação, especialmente na identificação das partes envolvidas em um processo ou identificação de penas em decisões judiciais.

Este trabalho relata alguns dos desafios iniciais enfrentados no desenvolvimento de um recurso léxico-computacional multilíngue baseado em *frames*. Nessa primeira etapa, aplicou-se uma metodologia de expansão dos *frames* da FrameNet para a criação dos *frames* do Direito brasileiro e busca dos equivalentes em inglês para os termos jurídicos brasileiros. Na metodologia de expansão, parte-se dos *frames* já existentes na FrameNet, adaptando os *frames* às peculiaridades do Direito brasileiro. Essa abordagem requer o conhecimento jurídico da área que se está descrevendo, pois nem sempre se encontra a correspondência no Direito brasileiro daqueles eventos jurídicos descritos pela FrameNet para o Direito americano. Nesta seção, descreve-se com mais profundidade a metodologia de expansão e os resultados obtidos a partir da análise contrastiva do *frame* *Criminal_process*.

2.1. Metodologia de expansão multilíngue de *frames*

Uma vez que a FrameNet possui uma ampla gama de *frames* semânticos capazes de descrever o significado de uma quantidade considerável de unidades lexicais, a metodologia de expansão de *frames* para diferentes línguas pode dar resultados frutíferos em um espaço reduzido de tempo, ao mesmo tempo que se aproveita todo o conhecimento já produzido pela FrameNet. Com frequência, projetos que estão desenvolvendo recursos lexicais baseados em *frames* para línguas além do inglês têm adotado essa metodologia.

A metodologia de expansão de *frames* para diferentes línguas consiste, basicamente, em aproveitar os *frames*, tidos como o componente conceptual da base de dados, e substituir as unidades lexicais evocadoras de *frame*, a parte da base de dados que é específica de cada

língua. Nesse processo, sempre que necessário, são feitos ajustes nos *frames* para representar as especificidades de cada língua.

Segundo Lönneker-Rodman (2007), quatro tipos de divergências, entre a base de dados da FrameNet e a base de dados de uma nova língua, podem ocorrer na construção de *framenets* em diferentes línguas por meio da metodologia de expansão. Essas divergências ocorrem entre *frames* semânticos, elementos de *frame*, tipos semânticos e relações entre *frames*.

Frame Semântico: segundo Lönneker-Rodman (2007), há pelo menos duas possibilidades para a criação de um novo *frame* em uma *framenet* de outra língua. Uma delas é a inadequação das descrições de *frame* da FrameNet, quando comparadas com o *frame* de outra língua. Para exemplificar a divergência entre *frames* em diferentes línguas, pode-se citar o *frame* *Notification_of_charges*, em inglês. Ao se compararem as unidades lexicais evocadoras do *frame* *Notification_of_charges*, como *acuse.v* (*acusar*), *charge.v* (*acusar*) ou *indict.v* (*pronunciar*), com seus equivalentes em português, percebe-se que as unidades lexicais *acusar* e *pronunciar* evocam dois *frames* diferentes. A unidade lexical *acusar* evoca o *frame* Denúncia e a unidade lexical *pronunciar* evoca o *frame* Pronúncia.

Elementos de Frame: a necessidade de criação de novos elementos de *frame* para comportar as diferenças entre as línguas também implica a criação de um novo *frame*. Os *frames* *Court_examination*, em inglês, e *Instrução*, em português, podem exemplificar a divergência de elementos de *frames* entre *frames* de línguas diferentes. O *frame* *Instrução*, apesar de ser comparável ao *frame* *Court_examination* na etapa do processo penal que ambos os *frames* representam, apresentou em português a necessidade de divisão do elemento de *frame* WITNESS em dois elementos de *frame*: TESTEMUNHA e RÉU.

Tipos Semânticos e Nuclearidade de Elementos de Frames: o *frame* da FrameNet e o *frame* de uma *framenet* de outra língua não são considerados equivalentes quando os elementos de *frame* nucleares na FrameNet não são nucleares no *frame* de uma *framenet* de outra língua, ou quando os elementos de *frame* apresentam um tipo semântico diferente.

Relações entre Frames: sempre que um *frame* apresentar um número diferente de elementos de *frame*, que esses elementos de *frame* apresentarem um tipo semântico diferente ou que a nuclearidade dos elementos de *frame* não for a mesma, as relações entre os elementos de *frame* mudarão. Qualquer uma dessas mudanças na outra língua acarretará a criação de um novo *frame* na nova *framenet*.

Como Boas (2005) aponta, a polissemia, os diferentes padrões valenciais, os diferentes padrões de lexicalização e as relações de paráfrase e de equivalência de tradução são desafios a serem enfrentados na criação de recursos léxico-computacionais multilíngues. No entanto, o que se pretende enfatizar na seção seguinte é que, no desenvolvimento de *framenets* para domínios especializados socialmente construídos⁶, como o Direito, o componente conceptual

⁶ Está-se chamando de domínios especializados socialmente construídos as áreas de conhecimento inteiramente criadas pela sociedade, como o Direito. Segundo Mattila (2006, p.105), “o direito não existe no mundo físico. Por ser inteiramente criado pelos humanos, o direito é sempre relacionado à cultura de uma determinada sociedade: portanto, constitui fenômeno cultural” (“Law does not exist in the physical world. Since it is entirely created by humans, law is always linked to the culture of any particular society: it therefore constitutes a social phenomenon.”). Talvez, dessas áreas de conhecimento criadas pela sociedade, o Direito seja a mais particular de todas, por apresentar um conjunto de normas válidas apenas para a sociedade que a criou. Outro exemplo de domínio especializado socialmente construído seriam as unidades de medidas, como *polegada*, *pé*, *centímetro* ou *metro*. As unidades de medidas variam, em alguns casos, de país para país. Assim, os EUA adotam unidades de medidas como *polegada*, *pé* e *milha*. O Brasil adota unidades diferentes: *centímetro*, *metro*, *quilômetro*.

pode ser fortemente influenciado por questões sociais. Na seção seguinte, são apontados alguns exemplos que demonstram diferentes graus de correspondência entre *frames* jurídicos em diferentes culturas jurídicas.

2.2. Etapas de um processo penal nos EUA e no Brasil e diferenças entre *frames*

O *frame* *Criminal_process* é um *frame* não-lexical (RUPPENHOFER ET AL. 2010). Isso significa que a FrameNet não apresenta unidades lexicais para esse *frame*. Ele representa um evento complexo, composto de diferentes etapas. Assim, o *frame* *Criminal_process* descreve as diferentes etapas de um processo penal de acordo com o sistema jurídico norte-americano. A Figura 1 mostra o *frame* *Criminal_process* e as relações que ligam os *subframes* entre si.

Em casos de *frames* complexos, como o *Criminal_process*, cada sequência de eventos ou estados é descrita como um *frame* relacionado ao *frame* complexo por meio da relação *Subframe* e aos demais *subframes* por meio da relação *Precede*. O *frame* *Criminal_process* é dividido em quatro *subframes*: Arrest, Arraignment, Trial e Sentencing. O *frame* Arraignment é dividido em três *subframes*: Notification_of_charges, Entering_a_plea, e Bail_decision. O *frame* Trial também apresenta três *subframes*: Court_examination, Jury_deliberation e Verdict.

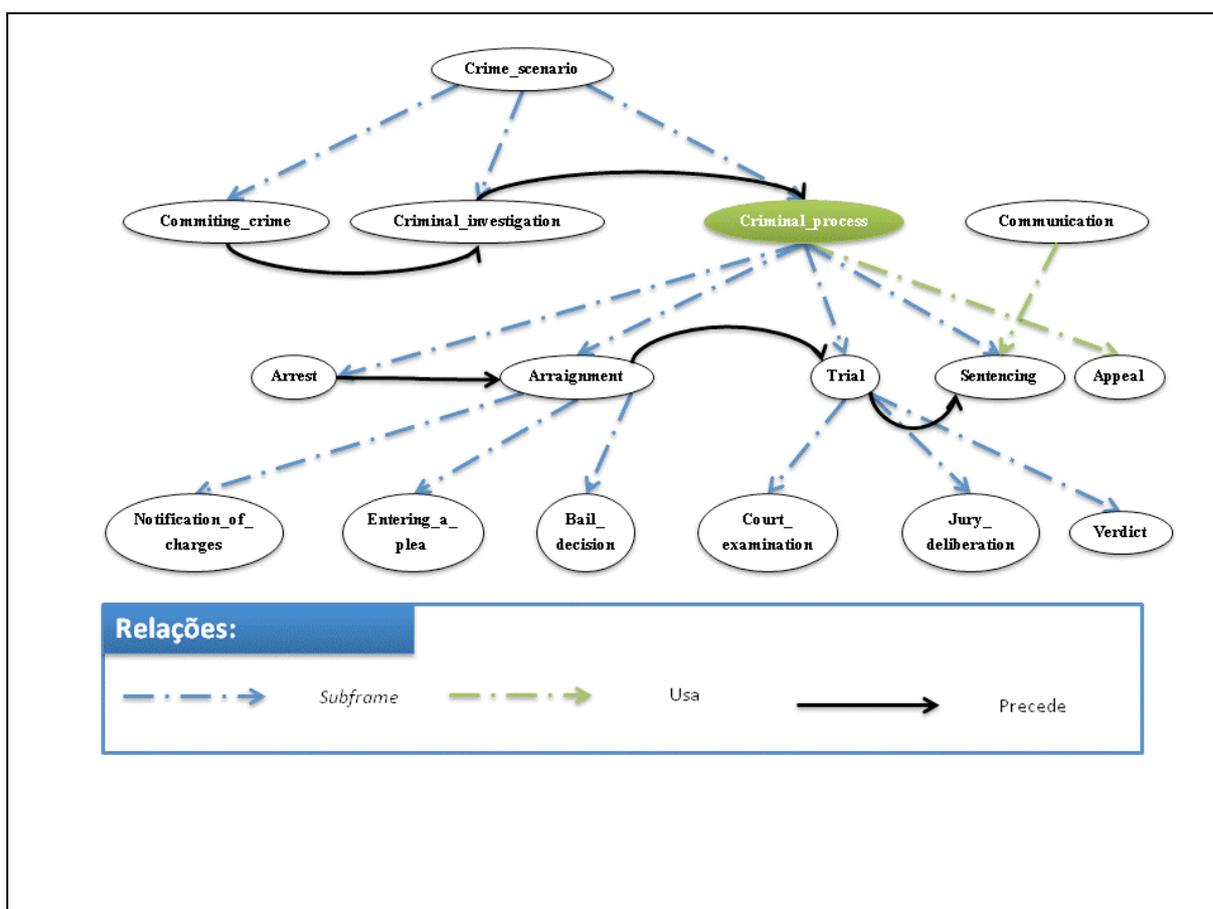


Figura 1: O *frame* *Criminal_process*

A partir da identificação de todos os *subframes* que compõem o *frame* *Criminal_process*, iniciou-se a análise contrastiva de todas as unidades lexicais evocadoras de *frames*. Essa análise foi feita de *frame* a *frame*. Primeiramente, com a ajuda de um dicionário jurídico bilíngue, foram identificadas unidades lexicais equivalentes em português às unidades lexicais em inglês. Quando as unidades lexicais em inglês podiam ser associadas a um equivalente em português, verificava-se se o conhecimento especializado evocado pelo equivalente em português era correspondente ao conhecimento especializado evocado pela unidade lexical em inglês⁷.

Para exemplificar a análise contrastiva das unidades lexicais, pode-se citar a unidade lexical *arrest.v*, evocadora do *frame* *Arrest* (Prisão). O *Noronha Dicionário Jurídico* (GOYOS JR., 1992) traz os verbos *prender*, *deter* e *capturar* como equivalentes para o verbo *arrest*. Então, o *frame* evocado pela unidade lexical em inglês é comparado ao *frame* evocado pelos equivalentes em português. A unidade lexical *arrest* evoca o *frame* *Arrest*, *frame* que descreve o evento da prisão, início de um processo penal nos Estados Unidos. A FrameNet define o *frame* *Arrest* como: “AUTORIDADES acusam um SUSPEITO, que está sob suspeita de ter cometido um crime e levam-no sob custódia”.⁸ Este *frame* traz como elementos de *frame* nucleares AUTORIDADES, ACUSAÇÕES, OFENSA e SUSPEITO.

Tanto *prender*, quanto *deter* e *capturar* evocam, segundo o cenário jurídico-criminal brasileiro, um *frame* de prisão comparável ao *frame* descrito pela FrameNet. Assim, pode-se perceber que, além da equivalência das unidades lexicais, o *frame* evocado pelas unidades lexicais em inglês é comparável ao *frame* evocado pelos seus equivalentes em português. No entanto, diferentemente do processo penal americano, que se inicia pela prisão, o processo penal brasileiro pode transcorrer sem a prisão do suspeito. Por esse motivo, o *frame* *Processo_penal*, criado de acordo com as etapas de um processo penal no Brasil, não se inicia pelo *frame* *Prisão*.

O *frame* *Arraignment* sucede o *frame* *Arrest*. O *arraignment* é uma etapa do processo penal americano que não encontra correspondente no processo penal brasileiro. Nessa audiência, o réu ouve a leitura das acusações pelas quais ele irá responder em juízo e declara a sua culpa ou a sua inocência. A FrameNet define o *frame* *Arraignment* como: “Em uma audiência no tribunal, um RÉU é informado das ACUSAÇÕES contra ele, (geralmente) dá sua RESPOSTA de inocência ou culpa e o JUIZ decide o montante da fiança, caso haja fiança”.⁹ O *frame* *Arraignment* possui como elementos de *frame* nucleares ACUSAÇÕES, RÉU e JUIZ.

Uma das unidades lexicais evocadoras do *frame* *Arraignment* é *arraignment.n*, que não apresenta equivalente no *Noronha Dicionário Jurídico*. Neste caso, a falta de correspondência conceptual, ou seja, a falta de um *frame* minimamente similar no processo penal brasileiro, reflete-se no nível lexical: não há como estabelecer um equivalente em português para o termo jurídico americano. No entanto, os *subframes* do *frame* *Arraignment* são evocados por unidades lexicais que apresentam equivalentes em português. Isso faz com que alguns desses *subframes* apresentem algum grau de correspondência com alguma etapa do processo penal brasileiro. Na seção seguinte,

⁷ É importante salientar que este trabalho foi realizado no contexto de um projeto multidisciplinar, envolvendo linguistas e advogados. Assim, todo o trabalho de estabelecimento de equivalências entre unidades lexicais e de construção de *frames* para o sistema jurídico brasileiro contou com a supervisão de especialistas da área jurídica.

⁸ “Authorities charge a Suspect, who is under suspicion of having committed a crime (the Charges), and take him/her into custody.”

⁹ “At a court hearing, a Defendant is informed of the Charges against him or her, (usually) enters a plea, and a decision is made by a Judge as to the amount of bail, if any.”

exemplifica-se esse caso de correspondência parcial por meio do *subframe* *Notification_of_charges*.

Essa não correspondência entre os *frames* que representam o processo penal americano e o processo penal brasileiro gera uma estrutura de *subframes* totalmente distinta entre o *frame* *Criminal_process* e o *frame* brasileiro *Processo_penal*. O resultado dessa análise contrastiva pode ser visto na Figura 2.

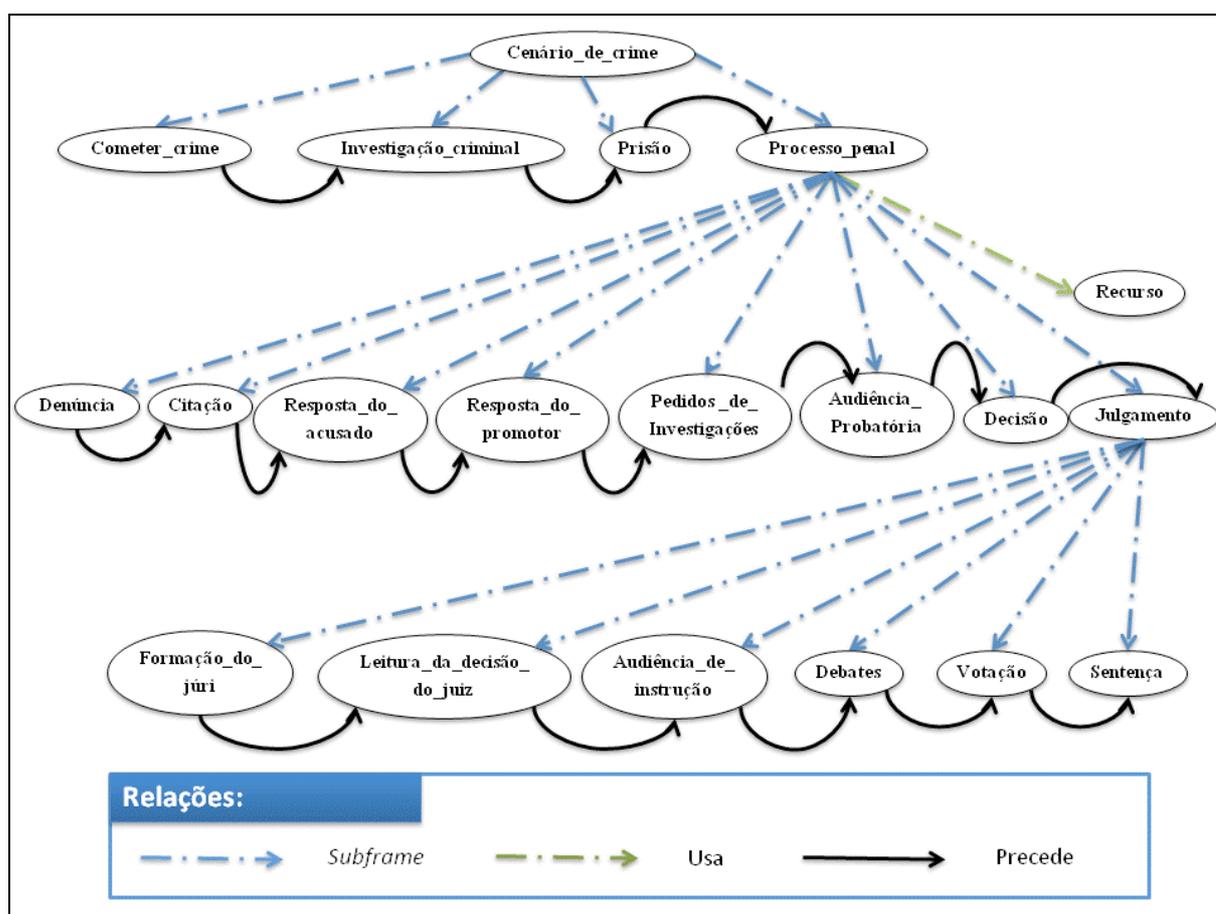


Figura 2: O *frame* *Processo_penal*

No caso da unidade lexical *deliberate.v*, uma das unidades lexicais evocadoras do *frame* *Jury_deliberation*, o *Noronha Dicionário Jurídico* apresenta como equivalente apenas o verbo *deliberar*. No entanto, no Brasil, os jurados são proibidos de falar entre si, podendo apenas votar a fim de decidir se o réu é inocente ou culpado. Assim, a palavra *deliberar*, apesar de ser equivalente de *deliberate*, não evoca um conhecimento especializado no processo penal brasileiro.

3. Correspondência de *frames*

A partir do estudo contrastivo das unidades lexicais evocadoras de *subframes* do *frame* *Criminal_process*, percebe-se que há casos em que a equivalência lexical existe, os dicionários apresentam equivalentes em português para a unidade lexical em inglês e os *corpora* de traduções apresentam termos jurídicos equivalentes em suas traduções. No entanto, o *frame* evocado pela unidade lexical em inglês não é correspondente ao *frame*

evocado pela unidade lexical em português. Esse é o caso das unidades lexicais que evocam o *frame* *Notification_of_charges*.

O *frame* *Notification_of_charges* descreve a primeira etapa do *frame* *Arraignment*. A FrameNet define o *frame* *Notification_of_charges* como: “o juiz ou outro oficial do tribunal (a AUTORIDADE), informa o ACUSADO das ACUSAÇÕES contra ele, por exemplo, as ações alegadas e as leis relevantes”.¹⁰ Segundo a FrameNet, o *frame* *Notification_of_charges* é evocado pelas unidades lexicais *accuse.v*, *charge.v*, *charge.n*, *indict.v* e *indictment.n*. A Tabela 1 mostra os equivalentes apresentados pelo *Noronha Dicionário Jurídico*. A consulta bidirecional demonstra imprecisão no estabelecimento de equivalentes em inglês e português.

Inglês	Português	Inglês
<i>accuse.v</i>	<i>acusar</i>	<i>to accuse; prosecute; charge; indict; incriminate; blame; arraign; renounce; acknowledge.</i>
	<i>denunciar</i>	<i>to denounce; accuse; inform against; report; proclaim.</i>
	<i>despesa</i>	
	<i>encargo</i>	
<i>charge.n</i>	<i>gravame</i>	
	<i>responsável</i>	
	<i>acusação</i>	<i>accusation; charge; incrimination; denunciation; prosecution; indictment.</i>
	<i>debitar</i>	
<i>charge.v</i>	<i>onerar</i>	
	<i>carregar</i>	
	<i>acusar</i>	
	<i>pronunciar</i>	<i>to indict; arraign.</i>
<i>indict.v</i>	<i>acusar</i>	
	<i>denunciar</i>	
<i>indictment.n</i>	<i>pronúncia</i>	<i>indictment; arraignment.</i>

Tabela 1: Unidades Lexicais Evocadoras do *Frame* *Notification_of_Charges*

Essa imprecisão em encontrar um equivalente apropriado pode estar relacionada à falta de correspondência entre as etapas do processo penal nos EUA e no Brasil. O verbo *to indict*, por exemplo, aponta como um equivalente o verbo português *acusar*. Quando se procura o equivalente de *acusar*, encontram-se, entre outros verbos, *to indict* e *to arraign*. No entanto, o *indictment* é uma fase anterior ao *arraignment*.

¹⁰ “The judge or other court officer (the Authority) informs the Accused of the Charges against him/her, i.e. the alleged actions and the relevant laws.”

O *Noronha Dicionário Jurídico* apresenta como equivalentes da unidade lexical *to indict* três unidades lexicais: *pronunciar*, *acusar* e *denunciar*. No entanto, essas unidades lexicais têm sentidos diferentes no contexto legal brasileiro. A acusação é algo mais genérico, qualquer pessoa pode acusar alguém. A denúncia é o documento pelo qual o Ministério Público dá início à ação penal; cabendo somente ao Ministério Público a denúncia. Já a pronúncia é a sentença preliminar apresentada por um juiz presidente do Tribunal do Júri para submeter um réu posteriormente a julgamento; ou seja, cabendo somente ao juiz pronunciar o réu. Esses três verbos são usados em etapas diferentes do processo penal no Brasil.

Apesar de as unidades lexicais do *frame* *Notification_of_charges* apresentarem equivalentes em português, o evento jurídico descrito pelo *frame* *Notification_of_charges* não possui correspondência completa com o processo penal brasileiro. Isso se deve, em primeiro lugar, ao fato de o *frame* *Notification_of_charges* ser *subframe* de um *frame* que não encontra correspondência no processo penal brasileiro. Em segundo lugar, as etapas do processo penal brasileiro que se sucedem da prisão do suspeito até o seu julgamento não são compatíveis com o processo penal americano.

Em geral, no Brasil, é apresentada denúncia contra o acusado e, após a denúncia, é realizada a audiência de instrução para coletar provas. A partir das provas coletadas na audiência de instrução, o juiz faz a pronúncia do acusado e, a partir deste momento, já como réu, ele vai a julgamento. Nos EUA, o acusado é levado ao juiz pela primeira vez para receber a fiança, comparecendo posteriormente para a audiência preliminar. Após a audiência preliminar, o processo é revisado por um grande júri, que faz a pronúncia do réu. Após o réu ser pronunciado pelo grande júri, ele deve comparecer ao *arraignment*, que é uma audiência em que o réu é informado das acusações pelas quais ele responde, tendo o réu que declarar a sua inocência ou culpa. Somente após o *arraignment*, o réu é levado a julgamento.

O *frame* *Notification_of_charges* necessita ser desdobrado em dois *frames* para representar corretamente o processo penal brasileiro: Denúncia e Pronúncia. No *frame* Denúncia, as unidades lexicais evocadoras de *frame* seriam *acusar.v*, *acusação.n*, *denunciar.v*, e *denúncia.n*. Já no *frame* Pronúncia, as unidades lexicais evocadoras de *frame* seriam *pronúncia.n* e *pronunciar.v*.

4. Implicações para o desenvolvimento de recursos lexicais multilíngues

Este trabalho tratou até aqui de problemas relacionados à equivalência de unidades lexicais em inglês evocadoras de *frames* jurídicos, tais como *arraignment*, *charge* e *indictment*, e da forma como essas questões de equivalência repercutiram na construção de um *frame* para o processo penal brasileiro. Cabe, neste momento, com base nas observações feitas até então, refletir sobre como questões de equivalência lexical e correspondência de *frames* podem interferir no desenvolvimento de recursos léxico-computacionais multilíngues para o domínio jurídico.

4.1. Desenvolvimento de recursos lexicais multilíngues para o domínio jurídico

Trabalhos de construção automática de recursos lexicais (PADÓ; LAPATA, 2005; PADÓ, 2007) afirmam que os métodos automatizados podem ser aplicados com sucesso ao desenvolvimento de recursos lexicais multilíngues baseados em *frames*. Com essa técnica, seria necessário apenas anotar um *corpus* em inglês com as etiquetas semânticas da FrameNet e transferir a anotação do *corpus* em inglês para um *corpus* paralelo em outra língua. Assim,

esses métodos estão fundamentados basicamente no estabelecimento de equivalências entre unidades lexicais de línguas diferentes. Nesta seção, discute-se como esses métodos automáticos podem falhar quando aplicados ao domínio jurídico, já que o Direito é uma área do conhecimento muito particular de cada sociedade. Aqui se apoiará esse argumento com exemplos do *frame* *Criminal_process*.

Como se percebeu no estudo do *frame* *Criminal_process*, vários *subframes* apresentam falta de correspondência com o sistema jurídico brasileiro em diversos níveis. Na FrameNet, o *frame* *Arraignment* descreve uma audiência em que o réu é informado das acusações contra si e, ao ser perguntado sobre sua culpa, ele pode se considerar culpado ou inocente. Um método automático de criação de léxicos provavelmente falharia neste ponto, pois o evento descrito pela FrameNet não encontra correspondente em português. Considerando-se um método baseado em *corpora* paralelos e em anotação automática, é preciso lembrar que se trata de *corpora* de textos traduzidos. Assim sendo, mesmo que a anotação do *corpus* em inglês fosse transposta para o *corpus* em português, o *frame* *Arraignment* não seria um *frame* representativo da realidade jurídica brasileira. A anotação do *corpus* em português estaria captando apenas a tradução para língua portuguesa de um evento que não existe no Brasil. Segue-se em (1a-b) e (2a-b) um exemplo de busca pela palavra *arraignment* em um *corpus* paralelo *online*.¹¹

- (1-a) Seven Ecuadoreans and two Colombians - one of whom is believed to be the leader of the gang - were arrested and brought before a judge for **arraignment**.
- (1-b) Sete equatorianos e dois colombianos - um dos quais seria o chefe da quadrilha - foram capturados e conduzidos a um juiz para a **formulação das acusações**.
- (2-a) At **arraignment**, the charge(s) is read to the defendant and he/she is permitted to enter a plea of guilty or not guilty.
- (2-b) Na **pronúncia**, é lida a acusação (ou acusações) ao arguido, que pode declarar-se culpado ou inocente.

No caso do *frame* *Notification_of_Charges*, embora unidades lexicais evocadoras deste *frame* apresentem equivalentes em português, esse *frame* necessitaria ser desdobrado em dois *subframes*: Denúncia e Pronúncia. Assim, de nada adiantaria tentar anotar sentenças em português com os elementos de *frame* apresentados pela FrameNet para o *frame* *Notification_of_charges*, pois seria necessária a criação de dois grupos diferentes de etiquetas para os *frames* Denúncia e Pronúncia. No exemplo abaixo (3a-b), também extraído do mesmo *corpus* paralelo, a unidade lexical *pronúncia* é traduzida como *indictment*, o que demonstra que uma pesquisa bidirecional a um *corpus* paralelo pode gerar resultados diversos.

- (3-a) O Presidente da República fica suspenso das suas funções a partir da data do trânsito em julgado do despacho de **pronúncia** ou equivalente e a sua condenação implica a destituição do cargo.

¹¹ <http://www.linguee.com.br/portugues-ingles/search?source=ingles&query=arraignment>

(3-b) The President of the Republic shall be suspended from his duties as of the date on which definitive **indictment** or its equivalent has been passed, and his conviction shall result in his discharge from office.

Nesses casos, a falta de correspondência entre os eventos jurídicos americanos e brasileiros poderia gerar resultados equivocados em uma transposição de anotação de um *corpus* em inglês para um *corpus* em português.

4.2. Anotação de *corpora* jurídicos com *frames* semânticos da FrameNet

As etiquetas semânticas da FrameNet têm sido utilizadas para anotação de textos em diferentes línguas, como o alemão (BURCHARDT ET AL., 2009) e o italiano (TONELLI; PIANTA, 2008), até mesmo para anotação de textos legislativos em italiano (VENTURI, 2011). No entanto, as divergências entre *frames* vistas até aqui chamam a atenção para a necessidade de se levar em conta possíveis diferenças entre sistemas jurídicos.

Dado que o Direito é uma área de conhecimento social e culturalmente construído, determinados *frames* podem ser específicos de uma cultura jurídica. Partir da equivalência das unidades lexicais para se anotar na língua alvo as unidades lexicais equivalentes com os mesmo *frames* evocados pelas unidades lexicais na língua fonte, como sugerem alguns trabalhos de transferência automática de anotação entre *corpora* paralelos (PADÓ; LAPATA, 2005; PADÓ, 2007), pode gerar alguns resultados não apropriados em se tratando de textos legais.

Os *corpora* jurídicos apresentam dois pontos de dificuldade: (a) a equivalência das unidades lexicais e (b) a equivalência dos conceitos jurídicos. Após identificar a unidade lexical a ser anotada, é necessário encontrar um equivalente em inglês para a unidade lexical que se quer anotar em português. Uma vez identificado um equivalente em inglês, é possível pesquisar a base de dados da FrameNet e verificar qual *frame* é evocado pela unidade lexical em inglês. A partir desse ponto, o anotador deverá analisar se o *frame* é apropriado para descrever o significado da unidade lexical em português.

Considerando a anotação da unidade lexical *denunciar*, a etapa inicial seria a busca de um equivalente em inglês. O anotador pode se sentir em dúvida ao procurar um equivalente apropriado em inglês. O equivalente mais óbvio para a unidade lexical *denunciar* seria *to denounce*. No entanto, se o anotador procurar pela unidade lexical *to denounce* na FrameNet, ele verá que o único *frame* evocado por essa unidade lexical é o *frame* Judgment_communication. Como esse não é um *frame* que descreve um evento relacionado ao domínio jurídico, o anotador terá que encontrar um outro equivalente para a unidade lexical *denunciar* e fazer uma nova busca na base de dados da FrameNet. Nesse caso, o anotador pode fazer uso de um dicionário jurídico bilíngue. O resultado possível de uma busca em um dicionário jurídico bilíngue pode ser uma entrada lexical como na Tabela 2, extraída do *Noronha Dicionário Jurídico* (GOYOS JR., 1992):

Português	Inglês
<i>denunciar</i>	<i>to denounce; accuse; inform against; report; proclaim.</i>

Tabela 2: Equivalentes de *Denunciar*

Nesse caso, a consulta ao dicionário pode tornar o trabalho do anotador ainda mais complicado. Em primeiro lugar, o verbo *to denounce* não evoca um cenário jurídico na FrameNet. Em segundo lugar, o verbo *to accuse* evoca um *frame* jurídico, o *frame* *Notification_of_charges*, mas, como já foi abordado anteriormente, esse *frame* não é apropriado para descrever o cenário jurídico evocado pela unidade lexical *denunciar* em português. Outro possível equivalente para *denunciar* poderia ser *inform against*, mas a FrameNet ainda não descreve unidades lexicais complexas de forma extensiva. Se o anotador procurar na base de dados da FrameNet por *inform*, os *frames* que ele encontrará não serão relacionados ao domínio jurídico. Esse *frames* são: *Reporting e Telling*.

Quando o anotador encontrar um equivalente para a unidade lexical a ser anotada, ele enfrentará o segundo desafio, que é a correspondência dos eventos jurídicos, pois equivalência lexical não é garantia de correspondência conceitual. Retomando o exemplo anterior, a unidade lexical *denunciar* tem como equivalente em inglês *to accuse*. No entanto, a unidade lexical *accuse* evoca o *frame* *Notification_of_charges*, enquanto o seu equivalente em português evoca o *frame* *Denúncia*, ou seja, o evento jurídico evocado pela unidade lexical em inglês é diferente do evento jurídico evocado por seu equivalente em português. Como o anotador está anotando um *corpus* jurídico, ele pode optar pelo *frame* *Notification_of_charges* para anotar a unidade lexical *denunciar*, porém os elementos de *frame* não corresponderão ao cenário jurídico evocado pela unidade lexical em português. Isso gera divergência no uso das etiquetas semânticas para anotação de textos em português.

Assim, em muitos casos, as relações entre *frames* da FrameNet não representam as etapas de um processo penal no Brasil. Portanto, o uso das etiquetas semânticas da FrameNet em domínios socialmente orientados, como o Direito, requer uma atenção especial para possíveis adaptações que sejam necessárias nos *frames*. Apesar de as unidades lexicais em português terem equivalentes em inglês, o cenário jurídico evocado pelas unidades lexicais em português é diferente do cenário jurídico evocado pelas unidades lexicais em inglês.

4.3. Correspondência de Elementos de *Frame*

Uma vez que um evento jurídico não apresenta correspondência, o *frame*, isto é, a representação esquemática de um evento jurídico, não apresentará correspondência entre dois sistemas jurídicos diferentes. Quando o cenário jurídico muda, os participantes daquele evento também podem mudar. De acordo com a definição da FrameNet para o *frame* *Notification_of_charges*, o juiz ou outra autoridade do tribunal (*ARRAIGN_AUTHORITY*) informa o acusado (*ACCUSED*) das acusações (*CHARGES*) que pesam contra ele. Os elementos de *frame* nucleares deste *frame* são *ARRAIGN_AUTHORITY*, *ACCUSED* e *CHARGES*.

Um anotador que esteja usando as etiquetas da FrameNet para anotar um *corpus* de textos jurídicos brasileiros terá que designar um *frame* semântico para cada unidade lexical evocadora de *frame* e elementos de *frame* para os sintagmas mais representativos de cada sentença. Ao se deparar com uma sentença cujo evocador de *frame* é a unidade lexical *acusar*, o anotador terá que identificar na FrameNet o *frame* apropriado. Como o anotador sabe que o equivalente em inglês para a unidade lexical *acusar* é *to accuse*, ele procura pela unidade lexical *to accuse* na base de dados da FrameNet e chega ao *frame* *Notification_of_charges*. O anotador perceberá imediatamente que o elemento de *frame* *ARRAIGN_AUTHORITY* não se ajusta bem ao cenário de um processo penal no Brasil. O que o anotador deve fazer neste caso? Ele deve considerar o elemento de *frame*

ARRAIGN_AUTHORITY apesar de esse elemento não se encaixar à descrição do processo penal brasileiro ou criar um novo *frame* para descrever o cenário jurídico brasileiro evocado pela unidade lexical *acusar*?

A falta de correspondência entre os eventos jurídicos nos procedimentos penais americano e brasileiro causa a falta de correspondência entre os elementos de *frame*. Assim, o anotador terá que decidir entre usar as etiquetas da FrameNet, ignorando a falta de correspondência de *frames* e elementos de *frame*, ou criar um novo *frame* cada vez que ele perceber a falta de correspondência dos eventos jurídicos e dos papéis desempenhados pelos participantes desses eventos nos sistemas jurídicos americano e brasileiro.

Considerações finais

O estudo contrastivo demonstrou que os *frames* semânticos em domínios socialmente construídos, como o Direito, apresentam um alto grau de divergência entre as línguas. A diferença entre os sistemas jurídicos norte-americano e brasileiro faz com que os eventos jurídicos não sejam os mesmos nos Estados Unidos e no Brasil. Isso provoca uma quebra entre a equivalência das unidades lexicais e a correspondência dos eventos jurídicos. Algumas unidades lexicais do inglês que apresentam equivalência em português podem não apresentar correspondência conceitual, ou seja, o significado é semelhante, mas o evento jurídico descrito por essa unidade lexical em inglês não é totalmente semelhante em português.

O tratamento da equivalência de *frames* é relativamente recente (LÖNNEKER-RODMAN, 2007). Em geral, os estudos de equivalência se detêm no estudo da equivalência de unidades lexicais. Os *frames* jurídicos são evidências da falta de equivalência de *frames* entre línguas, uma vez que esses *frames* representam um conhecimento socialmente construído, sendo, portanto, específicos de cada cultura e de cada país. A partir do estudo contrastivo dos *frames* semânticos da FrameNet e dos *frames* criados para o processo penal brasileiro, percebe-se que os *frames* jurídicos apresentam diferentes níveis de equivalência, variando desde a quantidade de elementos de *frame* até a natureza do evento jurídico descrito pelo *frame*.

A divergência entre os sistemas jurídicos norte-americano e brasileiro demonstrou a necessidade de criação de *frames* específicos para o processo penal brasileiro. A equivalência de *frames* entre as línguas é um fator a ser considerado quando se busca a precisão na anotação de *corpora* multilíngues com *frames* semânticos. No caso de anotação de *corpora* jurídicos, a falta de correspondência entre os sistemas jurídicos norte-americano e brasileiro traz à tona a necessidade de se criar um recurso lexical baseado em *frames* para a linguagem jurídica brasileira que possa ser utilizado para a anotação de *corpora*.

Agradecimentos:

Os autores agradecem o apoio de CAPES, CNPq, FAPERGS e FINEP.

Referências:

BERTOLDI, A.; CHISHMAN, R. L. O. Developing a Frame-Based Lexicon for the Brazilian Legal Language. In: *AI Approaches to the Complexity of Legal Systems - Models and Ethical Challenges for Legal Systems, Legal Language and Legal Ontologies, Argumentation and Software Agents*. Berlin / Heidelberg: Springer-Verlag, 2012. v. 7639. p. 256-270.

BOAS, H. C. Bilingual FrameNet Dictionaries for Machine Translation. In: *Proceedings of the Third International Conference on Language Resources and Evaluation*. Las Palmas, Spain. Vol. IV: 1364-1371. 2002.

BOAS, H. C. Semantic Frames as Interlingual Representations for Multilingual Lexical Databases. *International Journal of Lexicography*. 2005 18(4): 445-478.

BOAS, H. C. (Ed.) *Multilingual FrameNets in computational lexicography – methods and applications*. Berlin/New York: Mouton de Gruyter, 2009, 352p.

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.

CURTONI, P.; DINI, L.; DI TOMASO, V.; MOMMERS, L.; PETERS, W.; QUARESMA, P.; SCHWEIGHOFER, E.; TISCORNIA, D. Semantic access to multilingual legal information. In: SCHWEIGHOFER, E. (Ed.). *Proceedings of EU Info Workshop “Free EU Information on the Web: the future beyond the New Eur-Lex*. Brussels: Vrije Universiteit, 2005, p.1-11.

DINI, L.; PETERS, W.; LIEBWALD, D.; SCHWEIGHOFER, E.; MOMMERS, L.; VOERMANS, W.J.M.; Cross-lingual legal information retrieval using a WordNet architecture. In: *Proceedings of the 10th International Conference on Artificial Intelligence and Law, Bologna*. ACM Press: New York, 2005. p.163-167.

FELLBAUM, C. A semantic network of English: the mother of all wordnets. *Computers and the Humanities*. v.32, n.2-3, p.209-220, 1998.

FILLMORE, C. J. An alternative to checklist theories of meaning. In: *Proceedings of the first annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 1975, p.123-131.

FILLMORE, C. J. Scenes-and-frames semantics. In: ZAMPOLLI, A. (Ed.): *Linguistic Structures Processing: Fundamental Studies in Computer Science*, No. 59, North Holland Publishing, 1977, p.55-88.

FILLMORE, C. J. Frame semantics. In: THE LINGUISTIC SOCIETY OF KOREA (Eds.). *Linguistics in the Morning Calm*. Seoul: Hanshin, 1982, p.111-37.

FILLMORE, C. J. Frames and the semantics of understanding. *Quaderni di Semantica*. Vol.6, N.2, 1985, p.222-254.

FILLMORE, C. J.; JOHNSON, C. R.; PETRUCK, M. R. L. Background to FrameNet. *International Journal of Lexicography*. Vol.16, N°3, 2003, p.235-250.

GOFFMAN, E. *Frame Analysis*. New York: Harper, 1974. 596p.

GOYOS Jr., D. N. *Noronha's Legal Dictionary – Noronha Dicionário Jurídico: English-Portuguese, Portuguese-English – Inglês-Português, Português-Inglês*. 1.ed. São Paulo: Observador Legal, 1992.

LÖNNEKER-RODMAN, B. *Multilinguality and FramNet*. Technical Report. TR-07-001. Berkeley: ICSI, 2007.

MATTILA, H. *Comparative Legal Linguistics. Target*. Vol. 21. Hampshire: ASHGATE, 2006.

MILLER, G. A. WordNet: a lexical database for English. *Communications of the ACM*. New York: ACM Press. v.38, n.11, p.39-41, 1995.

MINSKY, M. *A framework for representing knowledge*. Artificial Intelligence Memo N°. 306. Cambridge, MA: Massachusetts Institute of Technology, 1974.

NUNES, M. G. V.; VIEIRA, F. M. C.; ZAVAGLIA, C.; SOSSOLOTE, C. R. C.; HERNANDEZ, J. A construção de um léxico para o português do Brasil: lições aprendidas e perspectivas. In: *Anais do II Encontro para O Processamento de Português Escrito e Falado*. Curitiba: CEFET-PR, p. 61-70.

OHARA, K. H.; FUJII, S.; SAITO, H.; ISHIZAKI, S.; OHORI, T.; SUZUKI, R. The Japanese FrameNet Project: A Preliminary Report. In: *Proceedings of Pacific Association For Computational Linguistics (PACLING'03)*. Halifax, Canada. August, 2003, p. 249-254.

PADÓ, S. *Cross-lingual Annotation Projection Models for Role-Semantic Information*. PhD Thesis. Saarbrücken: Universität des Saarlandes, 2007. 213p.

PADÓ, S.; LAPATA, M. Cross-lingual projection of role-semantic information. In: *Proceedings of HLT/EMNLP-05*. Vancouver: Association for Computational Linguistics, 2005, p.859-866.

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.

RUPPENHOFER, J.; ELLSWORTH, M.; PETRUCK, M. R. L.; JOHNSON, C. R.; SCHEFFCZYK, J. *FrameNet II: Extended Theory and Practice*. Berkeley: International Computer Science Institute, 2010. Disponível em: <http://framenet.icsi.berkeley.edu/>

SAGRI, M. T.; TISCORNIA, D. Semantic lexicons for accessing legal information. In: TRAUNMÜLLER, R. *Electronic Government*, LNCS3183. Springer-Verlag: Berlin/Heidelberg, 2004, p. 72-81.

SALOMÃO, M. M.M. FrameNet Brasil: um trabalho em progresso. *Calidoscópio*. Vol.7, No.3, 2009, p.171-182.

SCHMIDT, T. The Kicktionary – A multilingual lexical resource of football language. In: BOAS, H. C. (Ed.) *Multilingual FrameNets in computational lexicography: Methods and applications*. Berlin/New York: Mouton de Gruyter, 2009, p.102-132.

SNELL-HORNBY, M. *Tranlation studies: an integrated approach*. Amsterdam/Philadelphia: Benjamins Publishing Company, 1988, 163p.

SNELL-HORNBY, M. Of catfish and blue bananas: scenes-and-frames semantics as a contrastive ‘knowledge system’ for translation.’ In: DAM, V.; ENGBERG, J.; GERZYMISCH-ARBOGAST, H. (eds.) *Knowledge systems and translation*. Berlin/New York: Mouton de Gruyter, 2005, p. 193-206.

SUBIRATS, C.; PETRUCK, M. Surprise: Spanish FrameNet. *International Congress of Linguists. Workshop on Frame Semantics*, Prague (Czech Republic), July 2003. Disponível em: <http://gemini.uab.es/SFNpub/papers/subirats-petruck.pdf>. Acesso em: 15 julho de 2009.

TONELLI, S.; PIANTA, E. Frame Information Transfer from English to Italian. THE SIXTH INTERNATIONAL LANGUAGE RESOURCES AND EVALUATION (LREC'08). In: *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. Marrakech: European Language Resources Association (ELRA), p.28-30, 2008.

VANNEREM, M.; SNELL-HORNBY, M. Die Szene hinter dem Text: ‘scenes-and-frames semantics’ in der Übertzung. In: SNELL-HORNBY (ed.). *Übersetzungswissenschaft – Eine Neuorientierung. Zur Integrierung von Theorie und Praxis* (UTB 1415). Tübingen: Francke, 1986, p.184-205.

VENTURI, G. Semantic annotation of Italian legal texts: a frame-based approach. *Constructions and Frames*. Vol. 3, Nº. 1, 2011, pp. 46-79.

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.

VERMEER, H.; WITTE, H. *Mögen Sie Zistrosen? Scenes & frames & channels im translatorischen Handeln*. Heidelberg: Groos, 1990, 165p.

VOSSSEN, P. Introduction to EuroWordNet. *Computers and the Humanities*. v.32, n.2-3, p.73-89, 1998.

Received: 07/02/2013

Revised: 24/05/2013

Accepted: 28/06/2013

Published: 31/10/2013