

## Jur-FrameCorp: reflexões preliminares

Anderson Bertoldi

Programa de Pós-graduação em Linguística Aplicada – Universidade do Vale do Rio dos Sinos (UNISINOS)

andersonbertoldi@yahoo.com

**Resumo.** *Este trabalho apresenta o projeto Jur-FrameCorp. Trata-se de um projeto de doutorado que investiga a aplicabilidade da Semântica de Frames à construção de um léxico computacional jurídico para o Português do Brasil.*

**Abstract.** *This paper presents Jur-FrameCorp project. It is a doctoral research proposal that aims to investigate the applicability of Frame Semantics to the construction of a computational lexicon from Brazilian legal documents.*

**Palavras-chave:** semântica de *frames*; léxicos computacionais; terminologia jurídica

### 1. Introdução

O projeto “Jur-FrameCorp: Uma Base de Dados Lexical para a Recuperação de Informação Jurídica” insere-se no projeto FrameCorp<sup>1</sup> como um subprojeto, em nível de doutorado. Essa inserção ocorre à medida que ambos os projetos se preocuparão como o emprego da base de dados lexical FrameNet (Fontenelle, 2003) para o desenvolvimento de recursos lingüísticos para a Língua Portuguesa. Enquanto o projeto FrameCorp se encarregará da transposição das informações semânticas do FrameNet para um corpus, realizando assim a sua anotação semântica, o projeto Jur-FrameCorp se encarregará de transpor, manualmente, as entradas lexicais, relacionadas ao domínio jurídico, do Inglês para o Português.

“Jur-FrameCorp: Uma Base de Dados Lexical para a Recuperação de Informação Jurídica” é um projeto multidisciplinar a ser desenvolvido na área da Informática Jurídica (Aguiló-Regla, 2005). A Informática Jurídica é um campo de pesquisa essencialmente interdisciplinar que congrega pesquisadores das Ciências Jurídicas, Ciências da Computação e da Informação e da Linguística na busca por soluções em Tecnologia de Informação, tais como Web Semântica, Recuperação de Informação, Desenvolvimento de Softwares e Léxicos Jurídicos. A escolha do domínio jurídico para a realização desta pesquisa está relacionada: (i) ao potencial de inovação tecnológica da Informática Jurídica, (ii) à sua relevância social e (iii) a experiências anteriores nessa área (Bertoldi, 2007).

Um crescente interesse tem motivado a aplicação do paradigma FrameNet à construção de léxicos de diferentes línguas, tais como o German FrameNet (Boas, 2002;

---

<sup>1</sup> Projeto coordenado pela professora Dr<sup>a</sup>. Rove Chishman (UNISINOS).

2005), o Spanish FrameNet (Subirats e Petruck, 2003) e o Japanese FrameNet (Hasegawa et al., 2006). Estudos comparativos voltados à investigação da polissemia das línguas, dos padrões de lexicalização, dos padrões valenciais e das paráfrases e equivalentes de tradução, aplicando a Semântica de *Frames*, vêm sendo realizados em diferentes línguas. Este projeto vem somar esforços ao projeto FrameCorp nas discussões acerca da aplicabilidade da Semântica de *Frames* também para a Língua Portuguesa. A opção pelo uso do FrameNet como ponto de partida dá a este projeto uma dimensão aplicada, que são os resultados da pesquisa. Assim, ao final desta pesquisa, espera-se ter como produto tecnológico uma base de dados lexical do domínio jurídico em Língua Portuguesa.

## 2. Motivação e hipótese de pesquisa

O léxico computacional FrameNet, iniciado em 1997, tem desenvolvido uma base de dados lexical baseada na Semântica de *Frames* (Fillmore, 1982). A descrição do léxico da Língua Inglesa leva em consideração a valência sintática e semântica dos itens lexicais. A valência sintática especifica os tipos frasais (sintagma nominal, preposicional etc) e as funções gramaticais (sujeito, objeto etc). A valência semântica é descrita em termos de entidades que podem participar de um *frame* evocado por uma palavra, tais entidades são chamadas de “elementos *frame*” (Fillmore et al., 2003).

O projeto pioneiro inspirou a criação de “framenets” para outras línguas. Esse é o caso do German FrameNet (Boas, 2002), do Spanish FrameNet (Subirats e Petruck, 2003) e do Japanese FrameNet (Ohara, 2003). A discussão atual é a ligação dessas diferentes bases de dados, criando assim um léxico multilíngüe, utilizando o FrameNet como uma interlíngua (Boas, 2005). A criação de “framenets” e a sua conexão através de *frames* semânticos traz à tona questões como: (i) a polissemia das línguas, (ii) os diferentes padrões valenciais das línguas, (iii) os diferentes padrões de lexicalização das línguas e (iv) as relações de paráfrase e de equivalência de tradução (Boas, 2005).

O uso do FrameNet como uma interlíngua, ou seja, como um ponto de conexão entre diferentes bases de dados lexical, é um ponto de pesquisa em ascensão. Esse léxico computacional foi criado a partir de uma semântica pragmaticamente orientada, a Semântica de *Frames*, que leva em consideração os fatores situacionais e culturais envolvidos em um evento para assim descrever estrutura cognitiva desse evento, ou seja, é uma teoria a compor o grande espectro de teorias da chamada Linguística Cognitiva. O uso do FrameNet como interlíngua pressupõe que os *frames* semânticos sejam capazes de dar conta das diferentes línguas. Tendo em vista que a Semântica de *Frames* foca as questões sociais que estão envolvidas no conhecimento que um falante possui sobre sua língua e que os *frames* procuram representar a estrutura cognitiva desse conhecimento, coloca-se a questão de pesquisa envolvendo este trabalho: Os *frames* podem ser utilizados como interlíngua de um léxico especializado, como o do domínio jurídico?

Esta pesquisa parte da hipótese de que os *frames* não são capazes de suportar as diferenças semânticas das línguas, principalmente no que se refere a diferenças culturais, como as diferenças entre os sistemas jurídicos americano e brasileiro. Assim, o objetivo desta pesquisa é investigar em que medida diferenças léxico-conceituais (polissemia, os diferentes padrões valenciais e de lexicalização das línguas e as relações de paráfrase e de equivalência de tradução) e culturais (sistemas jurídicos diversos)

interferem no desenvolvimento de um léxico do domínio jurídico segundo o paradigma FrameNet. Uma experiência na construção de léxicos de domínio específico seguindo tal paradigma já foi realizada para a Língua Alemã. Trata-se do Kichtionary, um léxico multilíngüe para o futebol (<http://www.kicktionary.de/>).

### 3. Semântica de *Frames*, FrameNet e o domínio jurídico

“Jur-FrameCorp: Uma Base de Dados Lexical para a Recuperação de Informação Jurídica” é um projeto lingüístico-computacional que será desenvolvido segundo o modelo lingüístico cognitivista. A lingüística cognitiva é um paradigma científico relativamente novo, constituído há aproximadamente uma década e meia (Silva, 2004). A concepção central da lingüística cognitiva é a conceitualização humana do mundo, ou seja, que categorias os seres humanos criam para conceitualizar o mundo, qual a estrutura que esses conceitos tomam na mente humana, qual é a relação entre a linguagem, o mundo e a cognição humana.

A teoria cognitiva a ser investigada nesta pesquisa é a Semântica de *Frames* (Fillmore, 1982). Segundo Fillmore (1982), a Semântica de *Frames* é um programa de pesquisa em semântica empírica e uma estrutura descritiva para a representação dos resultados de tal pesquisa. A Semântica de *Frames* enfatiza as continuidades entre língua e experiência. Dessa forma, as palavras representam categorizações de experiência e essas categorias são estabelecidas através de situações motivadoras que têm como pano de fundo o conhecimento e a experiência (Fillmore, 1982). Fillmore (1982: 112) ainda afirma que a Semântica de *Frames* pode ser pensada como “um esforço para entender qual razão uma comunidade falante pode ter encontrado para criar a categoria representada pela palavra, e explicar o significado da palavra apresentando e esclarecendo essa razão”.

Fillmore (1982) concebe o *frame* como a caracterização de uma pequena “cena” ou “situação” abstrata, assim, para se entender a estrutura semântica de um verbo seria necessário entender as propriedades de tais cenas esquematizadas. “Pelo uso da palavra ‘*frame*’ para a maneira estruturada em que a cena é apresentada ou lembrada, nós queremos dizer que os *frames* estruturam o significado das palavras e que a palavra ‘evoca’ o *frame*” (Fillmore, 1982: 117).

A Semântica de *Frames* também é a teoria que está na base da construção do FrameNet (Fillmore et al., 2003). A base de dados lexical FrameNet, disponível on-line ([www.framenet.icsi.berkeley.edu](http://www.framenet.icsi.berkeley.edu)), contém mais de 10.000 Unidades Lexicais da Língua Inglesa, sendo mais de 6.100 totalmente anotadas, e um conjunto de mais de 940 *frames*. O FrameNet foi concebido inicialmente para ser um léxico monolíngüe, porém o desenvolvimento de projetos semelhantes para o Alemão (Boas, 2002; 2005), para o Espanhol (Subirats e Petruck, 2003) e Japonês (Hasegawa et al., 2006; Ohara et al., 2003) tem estimulado as discussões a respeito da possibilidade de uso do FrameNet como interlíngua. A possibilidade em discussão é o uso da base de dados de Língua Inglesa como ponto de conexão entre os diferentes *framenets*, criando, assim, um léxico multilíngüe (Boas, 2005). O uso dos *frames* semânticos como interlíngua pressupõe que eles estejam no nível cognitivo (conceitos) e que, portanto, possam ser utilizados como ponto de ligação do nível lexical (os itens lexicais das diferentes línguas), superando as dificuldades lexicais que seriam a falta de paralelismo entre as línguas.

A essa pressuposição se opõe a hipótese a ser testada nesta pesquisa. Esta pesquisa parte da hipótese de que os *frames*, por serem pragmaticamente motivados, não são os mesmos para as diferentes línguas, ou seja, não há um paralelismo conceitual entre as línguas, assim como não há um paralelismo lexical. Segundo Hirst (2004), mesmo quando as línguas lexicalizam conceitos semelhantes, os sentidos das palavras serão diferentes, pois cada língua lexicaliza diferentes perspectivas do mundo. Dessa forma, o objetivo desta pesquisa é investigar em que medida diferenças léxico-conceituais entre línguas, como a polissemia, os diferentes padrões valenciais e de lexicalização das línguas e as relações de paráfrase e de equivalência de tradução interferem no desenvolvimento de um léxico bilíngüe do domínio jurídico.

O caráter lingüístico-computacional desta pesquisa e sua delimitação teórico-metodológica faz com que ela siga a mesma metodologia empregada na construção de léxicos computacionais segundo o paradigma FrameNet: (i) estudos de um determinado *frame* semântico e dos itens lexicais da Língua Inglesa que já foram codificados pelo FrameNet, (ii) procura de equivalentes de tradução ou de paráfrases de tradução na língua-alvo para os itens da Língua Inglesa, (iii) estudo comparativo entre o Inglês e a língua-alvo propondo soluções para as dificuldades computacionais de transferências de dados lexicais de uma língua a outra. Para tanto, o método utilizado será a pesquisa empírica em *corpus* jurídico.

Estudos léxico-terminológicos têm mostrado que a polissemia das línguas é um desafio para tradução de textos jurídicos (Varó e Hughes, 2002). Enquanto os termos jurídicos estão no nível lexical, os *frames* do domínio jurídico estão no nível cognitivo. Assim, cabe a esta pesquisa analisar em que medida as diferenças lexicais se refletem no nível cognitivo, ou seja, em que medida as diferenças lexicais entre o Inglês e o Português resultarão em diferenças de *frames*, e assim dificultar o uso dos *frames* semânticos como interlíngua. É importante ressaltar o pressuposto cognitivista subjacente à teoria da Semântica de *Frames*: uma semântica pragmaticamente orientada que está relacionada à experiência e conhecimento de mundo de uma comunidade de falantes. Cabe aqui investigar se esse pressuposto cognitivista da Semântica de *Frames* é compatível com uma aplicação computacional que prevê a simples funcionalidade de uma base de dados multilíngüe. Caso tal pressuposto se chocar com a aplicação computacional, analisar as possibilidades de solução.

#### **4. Metodologia de pesquisa e etapas do trabalho**

A metodologia de pesquisa prevista para a execução deste projeto seguirá o paradigma FrameNet, que é a metodologia utilizada na construção dos diferentes *framenets* citados anteriormente. Isso permitirá uma futura integração das diferentes pesquisas. Essa metodologia pode ser resumida em quatro etapas: (i) estudos de um determinado *frame* semântico e dos itens lexicais descritos para esse *frame*, (ii) procura de equivalentes de tradução ou de paráfrases de tradução na língua-alvo para os itens da Língua Inglesa, (iii) estudo comparativo entre o Inglês e a língua-alvo, propondo soluções para as dificuldades computacionais de transferências de dados lexicais de uma língua a outra e (iv) edição de um léxico computacional para os itens lexicais da língua em estudo.

Nas subseções seguintes, apresenta-se o roteiro de trabalho e a forma como a metodologia descrita acima será empregada para a construção de uma base de dados lexical do domínio jurídico e para a investigação do uso do FrameNet como interlíngua.

#### **4.1. Constituição de um “framenet” jurídico**

A delimitação aplicada da pesquisa se dará a partir da seleção de um grupo de *frames* que irão compor um “framenet” jurídico. Dentre os 941 *frames* descritos pelo projeto FrameNet, foram selecionados, inicialmente, 46 *frames* ligados diretamente ao domínio jurídico, tais como, *Committing\_crime*, *Criminal\_investigation*, *Abusing*, *Law*. Esses 46 *frames* incluem informações como atividades jurídicas, tipos de ofensas e legislação. O próximo passo será selecionar um grupo menor de *frames* para iniciar o estudo.

#### **4.2. Construção de um *corpus* jurídico**

A construção do *corpus* jurídico é parte fundamental desta pesquisa, que é em seu cerne uma pesquisa empírica em Linguística Aplicada. Esta etapa da pesquisa inclui a definição da natureza dos textos a serem utilizados na composição do *corpus* (leis, doutrinas ou jurisprudências), extensão e representatividade do *corpus*, além da coleta dos textos em *sites* governamentais e preparação computacional dos textos (conversão de formato: *html* para *txt*).

#### **4.3. Análise lingüística**

A análise lingüística será baseada na Semântica de *Frames* (Fillmore, 1982) e seguirá a mesma metodologia adotada por projetos similares, tais como o German FrameNet (Boas, 2002), o Spanish FrameNet (Subirats e Petruck, 2003) e o Japanese FrameNet (Ohara et al., 2003). A metodologia consiste em utilizar a base de dados do FrameNet para comparar com a língua-alvo. Partindo-se dos *frames* jurídicos já selecionados, o trabalho com o *corpus* será de identificar os equivalentes de tradução para os itens lexicais do Inglês que evocam os *frames* jurídicos, tais como *investigate – investigar*, *appeal - apelar*. Ao realizar essa exploração será possível identificar em que medida as diferenças lexicais entre o Inglês e o Português podem dificultar: (i) a transposição dos dados do projeto FrameNet para o Português e (ii) o uso do FrameNet como interlíngua de uma base de dados bilíngüe ou multilíngüe.

#### **4.4. Edição da base de dados do domínio jurídico**

A construção de bases de dados lexical do domínio jurídico é uma necessidade para o avanço das pesquisas em Informática Jurídica. O projeto LOIS (Curtoni et al., 2005) criou uma base de dados lexical multilíngüe. O Jur-WorNet (Sagri et al., 2003) organizou um léxico jurídico da Língua Italiana e o conectou ao ItalWordNet (Roventini, 2003). Todos esses projetos demonstram a necessidade de maiores investimentos no processamento computacional do Português. A construção da base de dados lexical do domínio jurídico viria a suprir uma lacuna, que é a carência de recursos computacionais de domínio especializado para o Português. Para isso, será necessário o armazenamento das informações descritas pelo estudo teórico utilizando ferramentas próprias para a construção de léxicos computacionais. O editor de léxico a ser utilizado nesse trabalho, até o momento, está em aberto.

## 5. Reflexões preliminares

A pesquisa que está sendo descrita encontra-se em fase inicial. Nesta seção, apresentam-se algumas reflexões preliminares de futuros enfrentamentos relacionados à natureza terminológica desta pesquisa. O léxico FrameNet trata da língua geral, não apresentando preocupações com a criação de léxicos de domínio. Isso significa que a criação de um léxico jurídico a partir dessa base de dados deve examinar inclusive as relações de herança entre *frames* gerais e *frames* mais específicos (terminológicos). A subseção 5.1 examina alguns desafios a serem enfrentados na composição de um *corpus* jurídico e a subseção 5.2 traz algumas reflexões a respeito dos desafios de usar o FrameNet como léxico de partida para a construção de um léxico de domínio jurídico em Língua Portuguesa.

### 5.1. Desafios metodológicos: o *corpus*

Por tratar-se de um trabalho terminológico, o projeto Jur-frameCorp enfrenta o desafio de obter um *corpus* jurídico. Esse desafio se dá em dois níveis de profundidade: (i) A obtenção dos textos jurídicos e (ii) a representatividade do *corpus* para a finalidade proposta.

O primeiro enfrentamento do projeto Jur-FrameCorp será a obtenção de um *corpus* jurídico. De forma geral, os tribunais brasileiros não apresentam facilidade na disponibilização das suas bases de dados. Também as leis em formato digitalizado são difíceis de serem obtidas. Isso dificulta a construção de um *corpus* com uma grande cobertura em variedades e quantidade de textos. Tal dificuldade remete ao segundo desafio: a representatividade do *corpus*.

Considerando-se que se consiga obter uma grande quantidade de textos jurídicos, com uma variedade de gêneros satisfatória (leis, emendas, sentenças, acórdãos, pareceres), nada garante que esse *corpus* apresentará os equivalentes de tradução das unidades lexicais que estão presentes na base de dados do FrameNet. Até o presente momento, essa questão metodológica permanece em aberto.

### 5.2. Desafios teóricos: polissemia e especificidade de *frames*

A polissemia no FrameNet é tratada em termos da quantidade de *frames* que uma palavra evoca. Uma palavra como *appeal* (*apelar*), que evoca os *frames request* (*solicitação*) e *appeal* (*apelo*), possui duas Unidades Lexicais, em outras palavras, a quantidade de entradas, em diferentes *frames*, que uma mesma palavra pode ter. No caso da presente pesquisa, que tem por objetivo organizar um domínio jurídico para o FrameNet, palavras que deveriam ser consideradas polissêmicas, como *judge* (*julgar*) apresentam apenas uma entrada no FrameNet, e não no *frame* que interessa a esta pesquisa, ou seja, no *frame Judgment*.

A palavra *judge* está posta como evocadora do *frame Assessing* (*Avaliar*) e nenhuma ligação entre o *frame Assessing* e o *frame Judgment*. Isso significa que seguir a metodologia de expansão bilíngüe do FrameNet, tal qual realizada por outros projetos, ou seja, de escolher unidades lexicais na base de dados FrameNet e encontrar equivalentes na língua-alvo, não permitiria chegar à palavra *julgar* como evocadora do *frame Judgment*. Assim, tal metodologia deverá ser aplicada ao Português com o cuidado de tentar encontrar unidades lexicais ainda não analisadas pelo FrameNet. Cabe

ênfatizar que o FrameNet é um recurso de língua geral. Isso significa dizer que o trabalho que se pretende fazer seria uma extensão terminológica para esse recurso. O que se prevê até o momento, é subdivisão do *frame Judgment* em um *frame* mais geral, no sentido de “julgamento” como expressão de valores individuais, e uma ligação desse *frame* mais geral a um *frame* mais especializado, “julgamento” no sentido de ato jurídico, ou seja, uma especificação terminológica.

## 5. Referências

- AGUILÓ-REGLA, J. Introduction: Legal Informatics and the Conceptions of the Law. In: BENJAMINS V. R. et al. (Eds.). *Law and the Semantic Web: Legal ontologies, methodologies, legal information retrieval, and applications*, LNAI (3369). Berlin/Heidelberg: Springer-Verlag, 2005. p.18-24.
- BERTOLDI, A. *A semântica dos adjetivos: como e por que incluí-la em uma ontologia de domínio jurídico*. Dissertação (Mestrado em Lingüística Aplicada). UNISINOS, São Leopoldo. 2007.
- BENJAMINS, V. R. et al. Law and the semantic web: an introduction. In: BENJAMINS V. R. et al. (Eds.). *Law and the Semantic Web: Legal ontologies, methodologies, legal information retrieval, and applications*, LNAI (3369). Berlin/Heidelberg: Springer-Verlag, 2005. p.1-17.
- BOAS, H. C. “Bilingual FrameNet Dictionaries for Machine Translation”. In PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION. 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.
- CURTONI, P. et al. 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” OF JURIX 2005. Brussels: Vrije Universiteit. 2005, p.1-11
- BURCHARDT, A.; ERK, K.; FRANK, A.; KOWALSKI, A.; PADÓ, S. “SALTO -- A Versatile Multi-Level Annotation Tool”. Proceedings of LREC 2006, Genoa, Italy. 2006
- ERK, K.; KOWALSKI, A.; PADÓ, S.; PINKAL, M. “Towards a Resource for Lexical Semantics: A Large German Corpus with Extensive Semantic Annotation”. Proceedings of ACL 2003, Sapporo, 2003.
- FILLMORE, C. J.; JOHNSON, C. R.; PETRUCK, M. R. L. “Background to FrameNet”. *International Journal of Lexicography*. Vol. 16, n° 3, p. 235-250, 2003.
- FILLMORE, C. J. “Frame Semantics”. The Linguistic Society of Korea. *Linguistic in the Morning Calm*, Seoul, Hansinh Publishing Co., 1982, p.111-137.

FONTENELLE, T. A Bilingual Lexical Database for Frame Semantics. *International Journal of Lexicography*. Vol. 13, n° 4, p.232 – 248, 2000.

FONTENELLE, T. (ed.). *International Journal of Lexicography*. Special Issue: FrameNet and Frame Semantics. Vol. 16, n° 3, p. 231-385, 2003.

HASEGAWA, Y.; OHARA, K.; LEE-GOLDMAN, R.; FILLMORE, C. J. "Frame Interegion, Head Switching, and Translation: RISK in English and Japanese." ICCG4. FOURTH INTERNATIONAL CONFERENCE ON CONSTRUCTION GRAMMAR, Tokyo, Japan. September, 2006.

HIRST, Graeme. Ontology and the lexicon. In: Steffen Staab, S.; Studer, R. (Eds). *Handbook on Ontologies*. Springer: Berlin, 2004, p. 209-229.

LÖNNEKER-RODMAN, B. *Multilinguality and FrameNet*. ICSI Technical Report TR-07-001, Berkeley, CA, March 2007.

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. p.249-254. 2003.

ROVENTINI, A. et al. "Italwordnet: building a large semantic database for the automatic treatment of Italian". In: *Linguistica Computazionale*. Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali, ISSN. p.745-791, 2003.

SAGRI, M. T.; TISCORNIA, D.; BERTAGNA, F. Jur-WorNet. In: SOJKA, P. et al. (Eds.) SECOND INTERNATIONAL WORDNET CONFERENCE - GWC 2004. Brno: Masaryk University, 2003, p.305-310.

SILVA, Augusto Soares. Linguagem, Cultura e Cognição, ou a Linguística Cognitiva. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.), *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Coimbra: Almedina, vol. I, 2004, pp.1-18.

SUBIRATS, C.; PETRUCK, M. Surprise: Spanish FrameNet. *International Congress of Linguists. Workshop on Frame Semantics, Prague (Czech Republic), July 2003*.

SUBIRATS, C.; SAITO, H. 2004. "Spanish FrameNet and FrameSQL". 4TH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION. Lisbon (Portugal), May 2004.

VARÓ, Enrique Alcaraz; HUGHES, Brian. El español jurídico. Barcelona: Ariel, 2002.