

A ONTOGOV-MT: CONTRIBUIÇÕES LINGÜÍSTICO-TECNOLÓGICAS PARA A DEMOCRATIZAÇÃO DOS SERVIÇOS PÚBLICOS EM MATO GROSSO

Albano Dalla Pria¹
Gislaine Aparecida de Carvalho²

RESUMO

É cada vez maior a demanda de conhecimento lexical lingüisticamente sistematizado e formalmente representado para integrar sistemas de Processamento Automático de Línguas Naturais (PLN). Este trabalho, que se insere no âmbito de pesquisas interdisciplinares envolvendo Lingüística e PLN, propõe a investigação dos nominais e de suas relações semânticas, no português. Para tanto, propõe-se a análise de um conjunto de nominais coletados do portal e-MT. A proposta deverá fornecer uma representação lingüístico-computacional de relações semânticas dos nominais, no português. Ao sistematizar de modo formal um corpo de conhecimentos lingüísticos sobre nominais, esta proposta deverá contribuir para a construção de uma ontologia de domínio, a ONTOGOV-MT.

ABSTRACT

The development of NLP systems has created a great demand for formally represented and linguistically motivated lexical knowledge bases. This paper bridges part of this gap by proposing the investigation of the polysemy of nominals. Accordingly, the formal study of a set of nominals collected from the e-MT must reveal specific semantic relations that will contribute to a better understanding of Portuguese. Ultimately, the findings will be couched in terms of the ontology ONTOGOV-MT, which will lend itself straightforwardly to NLP programming.

Introdução

Este artigo constitui uma proposta projetiva que visa à representação lingüístico-computacional das relações semânticas e da polissemia dos nominais que deverá subsidiar a fase inicial de construção de uma ontologia específica para o governo do Estado de Mato Grosso.

Essa proposta insere-se no âmbito de pesquisas interdisciplinares envolvendo Lingüística e o Processamento Automático de Línguas Naturais. O PLN é uma subárea da Inteligência Artificial (IA) e visa à projeção e à implementação de sistemas computacionais para simulação do conhecimento lingüístico humano nesses sistemas (MENUZZI; OTHERO, 2005). Dada sua natureza teórica e aplicada, o PLN agrega diversas áreas do conhecimento, entre elas: a Lingüística, a Filosofia, a Psicologia, a Lógica, a Matemática, a Lingüística Computacional, a Ciência da Computação e a Inteligência Artificial. Cabe à Lingüística, nesse contexto, a sistematização e a formalização do conhecimento lingüístico necessário para a implementação dos sistemas. O PLN requer arcabouços gramaticais nos quais a competência e o desempenho lingüísticos de um falante nativo adulto sejam explicitados através de linguagens formais de representação da informação lingüística (PRIA, 2005).

Essa proposta é motivada (i) pela demanda cada vez maior de conhecimento lingüisticamente sistematizado e formalmente representado para a implementação de sistemas de Processamento Automático de Línguas Naturais (PLN) e (ii) pelo aprimoramento da interação entre usuários e sistemas computacionais por meio das línguas naturais (PYLYSHYN *et al.*, 1980). Nesse contexto, entende-se por sistemas de PLN programas computacionais “inteligentes” capazes de simular uma interação verbal através de uma língua natural (WINOGRAD, 1992).

A relevância da ONTOGOV-MT para o processamento computacional do português

A construção de bases de conhecimento léxico-semântico e lógico-conceitual formalmente representadas, como as ontologias e as *wordnets*³, embora demandem um trabalho alentado (porém

¹ Professor da Universidade do Estado de Mato Grosso – Campus Universitário de Alto Araguaia

² Professora da Universidade do Estado de Mato Grosso – Campus Universitário de Alto Araguaia

³ A *wordnet* é uma base relacional de conhecimento lingüístico na qual os itens lexicais estão interligados conceitualmente com base em relações sinonímia, e partir das quais outras relações (hiperonímia/hiponímia, meronímia/holonímia) são especificadas. Além da WordNet de Princeton

factível), porque requer pesquisa interdisciplinar e trabalho colaborativo, tem despertado grande interesse social, científico e tecnológico.

Do ponto de vista social, as ontologias facilitam a relação entre homens e sistemas computacionais empregados na WWW para oferecimento, por exemplo, de serviços governamentais *on-line*. Do ponto de vista lingüístico, uma ontologia é uma base relacional de conhecimento léxico-semântico e lógico-conceitual que possibilita investigações sobre as línguas naturais (EVENS, 1988). Do ponto de vista do PLN, uma ontologia configura um modelo abstrato formalmente representado do mundo que desejamos representar e é construída através da identificação de conceitos e de relações relevantes entre conceitos que podem ser processados em sistemas computacionais (FENSEL, 2000).

As ontologias projetadas no âmbito do PLN atendem às necessidades de reformulação da *World Wide Web* (WWW). A WWW foi inicialmente projetada, em 1991, como um serviço de gerenciamento de informação (BERNERS-LEE, 2000; BERNERS-LEE; HENDLER; LASSILA, 2001). Atualmente, dispõe de um volume de informações que aumenta vertiginosamente. Estima-se que estejam postados na rede mais de um bilhão de documentos (CHAVES; VIEIRA; RIGO, 2001). Se, por um lado, o elevado volume de informações publicadas na WWW representa o sucesso da proposta, por outro lado, dificulta o seu gerenciamento manual, dificultando ou comprometendo o acesso a tais informações. As dificuldades na busca de informações motivaram a reengenharia da concepção inicial da WWW, tal como a conhecemos hoje, com o intuito de facilitar, otimizar e agilizar o acesso ao repositório de informações nela disponibilizadas. Propôs-se a criação da *Semantic Web* (SW) (BERNERS-LEE; HENDLER; LASSILA, 2001).

O projeto SW prevê a implementação de ontologias ao projeto inicial da WWW, i. e., prevê a implementação de bases de conhecimento léxico-semântico e lógico-conceitual (EVENS, 1988) formalmente representado (GRUBER, 1993) sobre situações variadas do mundo real à WWW. Segundo Zavaglia et al (2005), “na Web, o uso de ontologias pode fornecer uma base de informações comum, bem como padronizada, englobando conceitos-chave que possam ser utilizados por serviços requisitados para cada situação particular. Em comércio eletrônico, por exemplo, o conjunto de informações oferecido pela ontologia pode ser utilizado para unificar e integrar definições de produtos oferecidos pelos mais variados pontos de venda, com um formato padrão e único. Além disso, as ontologias podem ser utilizadas em sistemas de recuperação da informação para melhorar a precisão e revocação dos documentos recuperados numa busca. Precisão é dada pela razão entre o número de respostas (documentos) corretas obtidas e o número total de respostas recuperadas. Revocação é dada pela razão entre o número de respostas corretas obtidas e o número total de respostas corretas possíveis” (p. 2).

A proposta de construção da ONTOGOV-MT pretende atender ao grande interesse social e político do Estado de Mato Grosso pela implementação de serviços governamentais disponíveis no portal *e-MT*. A ONTOGOV-MT não só poderá promover o aperfeiçoamento dos serviços oferecidos pelo portal, mas também a democratização desses serviços prestados pelo Estado de Mato Grosso, tão almejado por políticas governamentais e pela administração pública. Um projeto pioneiro, nesse sentido, é o Consórcio ONTOGOV⁴, financiado pela Alemanha, Grécia, Espanha e Suíça, cujo foco de investigação tem sido o enriquecimento de *e-Gov Services* com bases de conhecimento léxico-semântico.

As relações semânticas e polissêmicas pertinentes a uma base de conhecimento léxico-semântico

Desde a década de 1950, quando ocorreram as primeiras tentativas de simulação do processamento das línguas naturais em sistemas computacionais, um dos desafios para o PLN é a especificação explícita e formal de informações semânticas, visando-se a sua implementação nos sistemas⁵. Esse fato se deve a dois princípios básicos das línguas naturais que vem sendo abordados tanto no domínio lingüístico quanto do PLN: a ambigüidade e a polissemia (LEACOCK; RAVIN, 1999).

(Disponível em: <http://wordnet.princeton.edu/>), construída com base no inglês americano, e conhecida como a mãe das *wordnets*, o Consórcio *EuroWordnet* (VOSSEN, 1998) (Disponível em: <http://www.ilc.uva.nl/EuroWordNet/>) prevê a construção de *wordnets* para sete línguas européias: alemão, italiano, espanhol, belga, inglês, francês, tcheco e estônio. A *WordNet.PT* (MARRAFA, 2001) para o português europeu já se encontra *on-line* na WWW (Disponível em: <http://www.instituto-camoes.pt:8080/wordnet/>). Já a *WordNet-BR* para o português brasileiro encontra-se em fase de construção da sua base de conhecimento lingüístico (DIAS-DA-SILVA; OLIVEIRA; MORAES, 2002). A respeito da *WordNet*, consulte-se Fellbaum (1998).

⁴ Disponível em: <<http://www.ontogov.com/>>.

⁵ A tradução automática foi, ainda na década de 1950, uma das primeiras tentativas de simulação do processamento de línguas naturais em sistemas computacionais que vertiam textos e sentenças de uma

No domínio lingüístico, a semântica lexical tenta explicar a ambigüidade através da especificação das relações semânticas estabelecidas entre itens lexicais. Cruse (1986, 2000) define um conjunto de relações semânticas que podem ser identificadas entre itens lexicais. Citem-se, por exemplo, as relações de hiponímia/hiperonímia, meronímia/holonímia, sinonímia, antonímia, homonímia, agentividade e polissemia.

As relações de hiponímia/hiperonímia e meronímia/holonímia têm caráter determinante para a construção de bases de conhecimento léxico-semântico e lógico conceitual como ontologias e *wordnets* (FELLBAUM, 1998). A relação de hiponímia/hiperonímia, também conhecida por *é-um (tipo de)*, descreve uma relação de inclusão de significado de um termo (ou classe) mais específico no significado de um termo (ou classe) mais abrangente. Os hipônimos são termos mais específicos que dispõem de algumas, mas não todas, as propriedades de seus hiperônimos. Considerem-se os exemplos: (a) *A Casa Civil é um órgão estadual*, (b) *Blairo Maggi é um governador* e (c) *Secretário de estado de educação é um tipo de secretário*. *Casa Civil*, *Blairo Maggi* e *secretário de estado de educação* são hipônimos de *órgão estadual*, *governador* e *secretário*, respectivamente, e esses são hiperônimos daqueles.

A identificação dessas relações possibilita a especificação de inferências relevantes, por exemplo, para sistemas de busca de informações na WWW. Considerem-se os exemplos: (a) *A FAPEMAT é uma fundação*, (b) *As fundações são um tipo de órgão estadual*. Com base nas relações de hiponímia/hiperonímia estabelecidas em (a) e (b), pode-se derivar, por inferência, que *A FAPEMAT é um órgão estadual*.

A relação de meronímia/holonímia, conhecida como relação parte-todo, também tem se mostrado bastante produtiva nas línguas naturais. Os merônimos são termos cujo significado descreve partes constituintes do significado de outro termo, seu holônimo. Considerem-se os exemplos: (a) *Blairo Maggi é membro do governo*, (b) *O Secretário de estado de educação é membro da Secretária de Educação* e (c) *Os órgãos estaduais são parte do governo*. *Blairo Maggi*, *Secretário de estado de educação* e *órgãos estaduais* são merônimos de *governo*, *Secretária de Educação* e *órgãos estaduais*, respectivamente, e esses são holônimos daqueles.

As relações de meronímia/holonímia também possibilitam a especificação de inferências relevantes. Considerem-se os exemplos: (a) *Secretário de estado de educação é membro da Secretária de Educação*, (b) *Secretária de Educação é parte dos órgãos estaduais* e (c) *Os órgãos estaduais são parte do governo*. Com base nas relações de meronímia/holonímia estabelecidas em (a), (b) e (c), pode-se derivar, por inferência, que o *governo* tem uma *Secretária de Educação* e um *Secretário de estado de educação*.

O desafio de explicar a polissemia das línguas naturais tem motivado pesquisas que vão além da análise do comportamento de palavras ambíguas cujos significados distintos e não-relacionados são tradicionalmente denominados homônimos. A abordagem relacional descrita acima capta, por exemplo, dois sentidos distintos e não relacionados para o item *banco*, na frase *Hoje eu passei pelo banco*. Esses sentidos referem-se a *instituição financeira* ou *assento*. Entretanto, *banco* pode apresentar sentidos relacionados recuperados no contexto em que ocorrem. Considerem-se as frases: (a) *Este banco foi fundado em 1950*, (b) *Este banco está em greve* e (c) *O banco no qual depósito meu dinheiro fica ao lado do supermercado*. Nessas frases, *banco* refere-se a *instituição*, em (a), *pessoas*, em (b) e *instituição e edifício*, em (c). Os sentidos de *banco*, em (a), (b) e (c), estão todos relacionados a *instituição financeira*. A essa propriedade dos itens lexicais de apresentar significados relacionados, dependentes do contexto de ocorrência, Pustejovsky (1995) denominou polissemia complementar (ou lógica), em oposição à polissemia contrastiva (ou homonímia) que apenas distingue *banco₁* (= *instituição financeira*) de *banco₂* (= *assento*).

Aplicada ao domínio governo, a polissemia complementar descreve significados distintos, porém relacionados. Considerem-se as frases: (a) *O governo republicano foi fundado em 15 de novembro de 1889*, (b) *O governo está em greve*, (c) *O governo teve início em 1º de janeiro*, (d) *O governo tem duração de 4 anos* e (e) *O governo está em obras*. Os exemplos de (a) a (e) descrevem aspectos do significado de *governo*: em (a) denota uma instituição, (b) denota pessoas, (c) denota o processo de governar, (d) denota o evento de governar e (e) denota uma edificação.

Segundo Pustejovsky (1995), cada aspecto do significado de um item lexical refere-se a um “papel” relacionado ao significado desse item. O conjunto desses papéis é descrito formalmente em termos de uma Estutura Qualia que prevê (subespecifica) significados relacionados para um item lexical. Dada a complexidade do formalismo, serão descritos, neste artigo, apenas os papéis que o compõem e sua finalidade: o Papel Constituinte relaciona um objeto e suas partes constituintes (material, peso, partes e

elementos componenciais); o Papel Formal distingue os objetos dentro de um domínio maior (orientação, magnitude, forma, dimensionalidade, cor, posição), o Papel Télico especifica objetivo e função (objetivo de um agente na performance da ação, que especifica atividades) e o Papel Agente especifica fatores envolvidos na origem ou construção de um objeto (criador, artefato, cadeia causal).

Apesar das dificuldades já apontadas para o estudo do significado nas línguas naturais e para sua implementação em sistemas de PLN, as relações de significado descritas acima em termos de relações léxico-semânticas e da polissemia sistemática para alguns nominais apontam para a possibilidade de sistematização do conhecimento semântico dessa classe, visando-se à construção da ONTOGOV-MT.

Metodologia de trabalho para a construção da ONTOGOV-MT

Este artigo, enquanto projeção de construção da ONTOGOV-MT e, dada a natureza lingüístico-tecnológica dos estudos em PLN, orienta-se por uma metodologia que envolve o equacionamento dos problemas em três domínios: **lingüístico**, **lingüístico-computacional** e **computacional**, sendo que cada domínio envolve um conjunto próprio de atividades e competências. Cabe ao primeiro domínio a explicitação ou sistematização do conhecimento lingüístico; o segundo domínio é dedicado à representação formal do conhecimento lingüístico sistematizado; o terceiro domínio é dedicado à implementação computacional do conhecimento lingüístico sistematizado e formalmente representado anteriormente (DIAS-DA-SILVA, 1996).

Para a construção da ONTOGOV-MT, as ações deverão restringir-se aos domínios lingüístico e lingüístico-computacional. As ações do domínio lingüístico compreendem (i) a investigação do conhecimento lexical no processamento da linguagem e (ii) a investigação das relações semânticas e da polissemia sistemática dos nominais do domínio governo. As atividades do domínio lingüístico-computacional compreendem a formalização do conhecimento sistematizado no domínio lingüístico em termos de uma ontologia de domínio⁶.

Para atingir seus objetivos, esta proposta prevê o levantamento semi-automático de termos do domínio definido para investigação, i.e., parte do processo deve ser realizada por humanos e parte deve ser realizada automaticamente por ferramentas computacionais. Os termos que deverão compor a ontologia pretendida devem ser obtidos com base em um *cópus* de textos em língua portuguesa compilados a partir do Portal do Governo do Estado de Mato Grosso⁷ (*e-MT*). Para a compilação do *cópus*, devem ser adotados métodos manual e automático. Os textos devem ser coletados manualmente em formato eletrônico do portal *e-MT* e, posteriormente, submetidos ao *parser* PALAVRAS⁸, por exemplo. Essa ferramenta computacional, desenvolvida no Instituto de Linguagem e Comunicação da Universidade do Sul da Dinamarca, descreve a estrutura morfossintática das frases do português (BICK, 2000).

Para a extração dos termos relevantes a partir do *cópus* pode-se empregar o emprego do *plug-in OntoLT*⁹. Essa ferramenta, acoplada ao *parser* PALAVRAS, realiza a extração de termos candidatos no *cópus* anotado com base na frequência¹⁰ de estruturas definidas pelo usuário. Os resultados da extração são gerados em extensões compatíveis com editores de ontologias, facilitando a exportação de informação entre sistemas. A extração manual de termos candidatos a compor uma ontologia de domínio é um processo lento e demorado. A utilização do *parser* e do *plug-in* contribuem para a redução do tempo gasto nesse processo.

Extraídos os termos candidatos a compor a ONTOGOV-MT, devem ser analisados com base no referencial teórico da semântica lexical (CRUSE, 2000; EVENS, 1988) e da Teoria do Léxico Gerativo (TLG) (PUSTEJOVSKY, 1995), por exemplo. Esse referencial teórico também pode subsidiar a descrição manual das relações semânticas entre os nominais do domínio governo. A maneira como as relações semânticas estão descritas na *WordNet* (MILLER; FELLBAUM, 1991; FELLBAUM, 1998) também contribuem para o estudo das relações parte-todo (descrita pela dicotomia meronímia/holonímia) e para as relações de inclusão (descrita pela dicotomia hiperonímia/hiponímia) no âmbito da semântica lexical. A Estrutura Qualia, responsável não só pela descrição das relações parte-todo e de inclusão dos nominais, mas também pela descrição da polissemia complementar, contribuiria para a construção da ONTOGOV-MT.

⁶ Consulte-se Guarino (1998, 1997) sobre tipos de ontologia, dentre eles as ontologias de domínio.

⁷ Disponível em: <<http://www.mt.gov.br/>>.

⁸ Disponível em: <<http://visl.sdu.dk/visl/pt/parsing/nonautomatic/floresta.html?S=cetemcorpus#top>>.

⁹ Disponível em: <<http://olp.dfki.de/OntoLT/OntoLT.htm/>>.

¹⁰ A frequência pode ser um indício de que um termo é relevante no domínio em questão e que, portanto, pode ser um candidato a integrante da ontologia.

Por fim, as relações especificadas para o domínio governo a partir do *córpus* devem ser representadas formalmente. Para tanto, pode ser empregado, por exemplo, a ferramenta *Protégé*¹¹ (GENNARI et al, 2003). Esse editor de ontologias, desenvolvido na Universidade de Stanford, representa o conhecimento semântico hierarquicamente em termos de classes, atributos e relações e dispõe de recursos que convertem sua base de conhecimento em extensões (linguagem de máquina) que permitem sua implementação em diversas aplicações tecnológicas, tais como Sistemas de Perguntas e Respostas, serviços da *Web Semantica*, Pesquisa Semântica Multimídia e sistemas de busca de informações.

A ONTOGOV-MT: contribuições sociais, científicas e tecnológicas para o Estado de Mato Grosso

A proposta de construção da ONTOGOV-MT, i. e., de uma base de conhecimento semântico em forma de ontologia para o domínio governo inicialmente com base nos nominais extraídos do Portal do Estado de Mato Grosso, pode contribuir: (i) para o desenvolvimento de pesquisas lingüísticas e tecnológicas que visam à implementação e o processamento do português no portal *e-MT*; (ii) para a otimização da relação entre usuários e serviços oferecidos pelo portal *e-MT*; (iii) para a investigação de recursos tecnológicos que possam ser aplicados em pesquisas lingüísticas, tais como o *parser PALAVRAS*, o *plug-in OntoLT* e o editor *Protégé*; (iv) para o refinamento e reformulação de propostas teóricas na investigação lingüística, principalmente no que se refere aos nominais; (v) para a validação tanto de teorias quanto de metodologias empregadas no desenvolvimento do trabalho interdisciplinar envolvendo Lingüística e PLN; (vi) para o avanço dos conhecimentos sobre os nominais e sobre sua descrição, quer do ponto de vista estritamente lingüístico quer do ponto de vista do ensino de língua portuguesa; (vii) para o fortalecimento do trabalho interdisciplinar envolvendo Lingüística e PLN no Estado de Mato Grosso; e (viii) para a formação e o aperfeiçoamento profissional de pesquisadores no âmbito do PLN, ressaltando-se a importância do profissional de Letras para atuação nessa área.

Referências bibliográficas

- ARNOLD, D.J et al. *Machine Translation: an Introductory Guide*. London: Blackwells, 1994.
- BERNERS-LEE, T.; T.; HENDLER, J.; LASSILA, O. *The semantic web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*. New York: Scientific American, 2001. Disponível em: <<http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2>>. Acesso em 25 de set. de 2005.
- BICK, E. *The Parsing System PALAVRAS: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. PhD thesis, Arhus University, 2000.
- CHAVES, M. S.; VIEIRA, R.; RIGO, S. Uso de ontologias para gerenciamento e acesso a documentos na web. In: V Oficina de Inteligência Artificial, 2001, Pelotas. *Anais da V Oficina de Inteligência Artificial*. Pelotas: EDUCAT, 2001.
- CRUSE, A. *Lexical semantics*. Cambridge: Cambridge University Press, 1986.
- CRUSE, D. A.. *Meaning in language: an introduction to semantics and pragmatics*. New York: Oxford University Press, 2000.
- DIAS-DA-SILVA, B. C. *A face tecnológica dos estudos da linguagem: o processamento automático das línguas naturais*. 1996. 272f. Tese (Doutorado em Lingüística e Língua Portuguesa) – Faculdade de Ciências e Letras de Araraquara, Unesp, Araraquara, 1996.
- DIAS-DA-SILVA, B. C.; OLIVEIRA, M. F.; MORAES, H. R. Groundwork for the development of the Brazilian Portuguese Wordnet. In: *Advances in natural language processing*. Berlin: Springer-Verlag, 2002. p.189-196.
- EVENS, M. *Relational models of the lexicon*. Cambridge: Cambridge University Press, 1988.
- FELLBAUM, C. *WordNet: an electronic lexical database*. Cambridge: The MIT Press, 1998.
- FENSEL, D. *Ontologies: silver bullet for knowledge management and electronic commerce*. Berlin: Springer-Verlag, 2000. Disponível em: <<http://www.cs.vu.nl/~dieter>>.

¹¹ Disponível em: <<http://protege.stanford.edu/>>.

- GENNARI, J. H.; MUSEN, M. A.; FERGERSON, R. W.; GROSSO, W. E.; CRUBZY, M.; ERIKSSON, H.; Noy, N. F.; Tu, S. W. The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. *International Journal of Human-Computer Studies*, v. 58, n. 1, p. 89-123, 2003.
- GRUBER, T. R. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, v. 5, n. 2, p. 199-220, 1993.
- GUARINO, N. Formal Ontology and Information Systems. *Proceedings of FOIS'98*. TrentoAmsterdan: IOS Press, p. 3-15, 1998.
- GUARINO, N. Understanding, Building and Using Ontologies. *International Journal of Human and Computer Studies*, v. 46, p. 293-310, 1997. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>>.
- LEACOCK, C.; RAVIN, Y. (eds.) *Polysemy: Theoretical and Computational Approaches*. Oxford: Oxford University Press, 2000.
- MARRAFA, P. *WordNet do português: uma base de dados de conhecimento lingüístico*. Lisboa: Instituto Camões, 2001.
- MENUZZI, S. de M.; OTHERO, G de Av. *Lingüística computacional: teoria e prática*. São Paulo: Parábola, 2005.
- MILLER, G.; FELLBAUM, C. Semantic networks of English. *Cognition*, v. 41, p. 179-229, 1991.
- NIRENBURG, S. (org.) *Machine Translation*. Cambridge: Cambridge University Press, 1987.
- PRIA, A. D. *Uma proposta de representação lingüístico-computacional do comportamento sintático e semântico de adjetivos no sintagma nominal do inglês e do português*. Dissertação (Mestrado em Lingüística e Língua Portuguesa) – Faculdade de Ciências e Letras da Universidade Estadual Paulista, Araraquara, 2005.
- PUSTEJOVSKY, J. *The generative lexicon*. Cambridge: The MIT Press, 1995.
- PYLYSHYN, Z. W. et al. Understanding natural language. *IBM Journal of research and development*, v. 35, p. 1-20, 1980.
- VOSSSEN, P. (Ed.) *EuroWordNet: a multilingual database with lexical semantic networks*. Dordrecht: Kluwer, 1998.
- WINOGRAD, T. *Understanding natural language*. New York: Academic Press, 1972.
- ZAVAGLIA, C.; OLIVEIRA, L. H. M.; NUNES, M. G. V.; TELINE, M. F.; ALUÍSIO, S. M.. *Avaliação de Métodos de Extração Automática de Termos para a Construção de Ontologias*. Relatório Técnico do NILC. São Carlos, 2005.