

EX234 ESPECIFICAÇÃO DE TRAÇOS SEMÂNTICOS DOS ITENS LEXICAIS. Ariani Di Felippo (Bolsista FINEP); Lúcia Helena Machado Rino (O); (DC/UFSCar)

Este trabalho faz parte de um Projeto PADCTIII - Processo FINEP-Itautec/Philco No. 03-CE-01/98-01/02-10 e PIBIC/CNPq, intitulado "*Revisor Gramatical e Ferramentas de Auxílio à Escrita*" - mais especificamente, de um subprojeto denominado "*Especificação de traços semânticos dos itens lexicais*". Este projeto é desenvolvido no NILC - Núcleo Interinstitucional de Linguística Computacional/São Carlos, instituição esta que agrega pesquisadores da USP-São Carlos, UFSCar e UNESP-Araraquara. Esse subprojeto visa o aprimoramento do ReGra, o Revisor Gramatical da Língua Portuguesa desenvolvido pelo NILC e comercializado pela ITAUTEC/PHILCO desde 1995, no que diz respeito ao tratamento da ambigüidade categorial decorrente da ocorrência de homonímia.

Assim, seu objetivo é fornecer material lingüístico tanto para o detalhamento do léxico do ReGra, quanto para a incorporação de um módulo de processamento semântico à análise gramatical já realizada pelo ReGra. Vale notar que a versão atual do ReGra não inclui qualquer tipo de processamento semântico e, por essa razão, seus diagnósticos de revisão gramatical são, muitas vezes, problemáticos. Para tanto, utilizando a ferramenta de análise estatística WordSmith, buscamos no corpus do NILC (37 milhões de palavras), primeiramente, a distribuição de freqüência e a ocorrência no contexto de alguns verbos plenos de sentido que pudessem ser computacionalmente tratáveis.

Após a compilação dos contextos nos quais esses itens lexicais aparecem, utilizamos a Teoria de Valências para fazer um levantamento de suas propriedades sintático-semânticas, a fim de identificar seus traços argumentais mais ocorrentes. Essas informações devem dar origem à inclusão de informações semânticas ao léxico do ReGra, para que seja possível tratar os casos de homonímia durante o processamento sintático das sentenças sob revisão.

Esse trabalho lingüístico é de suma importância para o aprimoramento do ReGra, pois deve, além de permitir a especificação lingüística adequada, apontar as classes de problemas que sejam operacionalmente viáveis, quer pelo volume ou complexidade da especificação lingüística a ser incorporada ao revisor, quer pela complexidade computacional envolvida.

O referido trabalho está em andamento, sendo assim, no CIC, apresentaremos um caso (estrutura sintática) computacionalmente tratável cuja ambigüidade categorial decorrente da ocorrência de homonímia pode ser tratada de acordo com o método explicitado.