

242

**ESTRUTURAS SINTÁTICAS E EXTRAÇÃO DE INFORMAÇÃO.** *Éder Alves de Macedo, Renata Vieira (orient.)* (Programa Interdisciplinar de Pós-Graduação em Comp, Centro de Ciências Exatas, UNISINOS).

A busca pela construção de mecanismos de extração automática de informação em textos é um dos motivos pelo qual a Linguística Computacional tem estado em constante atuação. A fim de auxiliar na construção de Mapas Conceituais, esta pesquisa visa identificar padrões de estruturas sintáticas que apresentam informações com maior relevância semântica. Os métodos utilizados foram a análise de estruturas sintáticas com base nos padrões da Gramática da Língua. Foram selecionados textos jornalísticos que, em seguida, passaram pela análise sintática e pela extração das suas estruturas, efetuadas automaticamente. Meios de extração via sintaxe foram testados e, em seguida, analisaram-se as informações contidas nos extratos obtidos. Optou-se, primeiramente, pela extração dos seguintes componentes das sentenças: Sujeito, Verbo e Objeto Direto. A partir disso, visando o aprimoramento dessa extração, estudou-se a limitação do resultado obtido: observou-se que esta combinação não porta informações realmente representativas. Com isso, fez-se a análise de novas combinações com outros termos sintáticos para que a relevância dos termos fosse avaliada e comparada à informação contida nos textos. Viu-se que adjuntos adverbiais são nodos que representam circunstâncias na oração e, portanto, possuem grande importância segundo um dado contexto. Por outro lado, os apostos possuem conteúdos suplementares que estão ligados às informações externas no texto, isto é, ao conhecimento de mundo do leitor. De modo geral, inferiu-se que tanto os termos essenciais e integrantes (sujeitos, verbos, objetos) quanto os acessórios (adjuntos adverbiais, adjuntos adnominais, apostos) representam, em textos jornalísticos, áreas de informação relevante, pois são fontes complementares em relação ao contexto maior do discurso. Para a continuidade da pesquisa, o mesmo processo será efetuado em um corpus composto de textos didáticos, além da análise das relações de subordinação e coordenação entre as orações. (FAPERGS/IC).