

072

EXPLORANDO GRANDES EXTENSÕES TEXTUAIS PARA FINS TERMINOLÓGICOS.

Caroline Vieira Guimaraes dos Santos, Maria da Graca Krieger (orient.) (Departamento de Letras Clássicas e Vernáculas, Instituto de Letras, UFRGS).

O projeto TERMISUL prepara um glossário multilíngüe (português/inglês/alemão/francês/espanhol) na área de gestão ambiental. Como corpus textual eletrônico utilizamos, em um primeiro momento, os textos da Série ISO 14000 – Sistema de Gestão Ambiental (SGA) – e manuais de implantação desse sistema. Posteriormente, foi introduzida uma terceira categoria de textos que reflete uma linguagem empresarial referente à divulgação de produtos de organizações que já têm o SGA implantado e que buscam promover uma imagem de “empresa verde”. Com a inclusão dessa nova categoria de textos, a extensão do corpus aumentou consideravelmente, exigindo um tempo muito maior de leitura linear para a coleta de candidatos a termo. Por esse motivo, resolvemos recorrer a ferramentas computacionais, cuja utilização inicial não se revelou produtiva para uma coleta eficiente e rápida. Diante disso, desenvolvemos uma metodologia com base nos pressupostos teóricos da descrição linguística baseada em corpus, para que nos auxiliasse na busca de candidatos a termo dessa categoria. Nesse sentido, tentamos unir a abordagem linguística adotada e a sua aplicação informatizada. Esta comunicação tem como objetivo relatar os procedimentos implementados com a finalidade de aprimorar tais buscas e de evitar desperdício de tempo e de trabalho, sem perda de qualidade. Assim, apresentamos as diferentes etapas do trabalho e aplicações desenvolvidas com os recursos oferecidos pelo software WordSmith 2.0 (Scott, 1996), explicando como foram aproveitadas as suas potencialidades para a identificação de informações linguísticas relevantes para a pesquisa terminológica desenvolvida em um corpus de 150.000 palavras. (PIBIC/CNPq-UFRGS).