
- **LÉXICO E INFORMÁTICA: UMA RELAÇÃO NECESSÁRIA**
Coordenador(a): *Gladis Maria de Barcellos Almeida*

Há algumas décadas estamos acompanhando o crescimento de estudos e pesquisas na área de Linguística Computacional e o conseqüente aprimoramento de ferramentas computacionais voltadas para o processamento de língua natural (PLN). Entretanto, no que se refere ao tratamento do léxico, geral ou especializado, observa-se um tímido impacto dessas ferramentas na pesquisa lingüística, sobretudo no Brasil. Pesquisadores brasileiros ainda se vêem obrigados a utilizar ferramentas não desenvolvidas especificamente para tarefas nas quais elas estão sendo empregadas, ou ainda tendo que desenvolver determinadas etapas metodológicas de suas pesquisas sem o auxílio da informática. Nosso objetivo nesse simpósio é promover uma ampla discussão sobre a relação Léxico e Informática, dando relevo para as seguintes questões: como a informática está interferindo na prática lexicológica/terminológica? Que ferramentas são utilizadas pelos diversos grupos? Quais os problemas enfrentados quando se inserem etapas semi-automatizadas à metodologia de trabalho? Quais procedimentos a serem adotados para tornar a pesquisa menos dispendiosa no que se refere a tempo? No diálogo interdisciplinar com a Informática, que se tem revelado bastante proficuo em algumas universidades brasileiras, o que pedir para os informatas, ou dito de outro modo: quais são as nossas demandas?

A PRÁTICA TERMINOLÓGICA E A INFORMÁTICA

Gladis Maria de Barcellos Almeida (UFSCAR)

O uso de terminologias sistematizadas contribui para tornar mais eficaz a comunicação entre especialistas. Entretanto, para se empreender a tarefa de sistematizar terminologias de forma ágil, é fundamental que haja ferramentas computacionais compatíveis com esse tipo de empreendimento. A Terminologia e a Informática são duas áreas que estão integradas, há pelo menos três décadas, em países desenvolvidos, com o objetivo de gerar produtos terminológicos melhores e mais fiáveis. É fato que, no Brasil, tal realidade vai se dar muito tardiamente e, ainda assim, os poucos terminólogos brasileiros que desenvolvem produtos terminológicos operam na maioria das vezes com ferramentas adaptadas e não elaboradas especificamente para a finalidade das suas pesquisas, o que mostra a grande necessidade de a Terminologia aliar-se cada vez mais à Informática. Nossa intenção aqui é fazer um breve relato de nossa experiência de elaboração de dicionários terminológicos, dando relevo não só para os problemas enfrentados no tocante a determinados procedimentos metodológicos como também para as soluções buscadas para a minimização desses problemas.

ELABORAÇÃO DE UM DICIONÁRIO DE VERBOS PORTUGUÊS-ALEMÃO

João Moraes Pinto Junior (UNESP)

A forma de elaboração de dicionários passou por uma mudança radical a partir do momento em que o computador começa a ser empregado como uma ferramenta na montagem de verbetes. Tem-se a impressão de que o computador foi desenvolvido especialmente para trabalhar com as palavras. Houve o que se pode denominar de uma aliança perfeita. O problema que se coloca atualmente não é quanto ao uso do computador para a digitação de verbetes ou mesmo para a busca aleatória de abonações para se compor a nomenclatura de um dicionário, mas o de utilizar ferramentas apropriadas para que, a partir de um corpus, se consiga elaborar os dicionários de forma prática e objetiva. Se por um lado sabemos quais são os pré-requisitos para a formação de um corpus computadorizado e quais os elementos teóricos que devem nortear um dicionário de usos, precisamos agora de ferramentas computadorizadas, altamente desenvolvidas, para que possamos elaborar dicionários monolíngües ou bilingües. O processo na elaboração do dicionário de verbos Português-Alemão nos mostrou que há necessidade de uma interação entre o lexicógrafo e o informata de modo que, cada um dentro da sua área de atuação, consiga solucionar os problemas que surgem no decorrer da pesquisa e que necessitam de uma solução especializada.

INFORMATIZAÇÃO E RECONHECIMENTO DE PERFIS DA LINGUAGEM DE QUÍMICA EM TEXTOS: PROBLEMAS E VANTAGENS DO RECONHECIMENTO DE PALAVRAS EM GRANDE ESCALA

Maria José Bocorny Finatto (UFRGS), Susana de Azeredo (UFRGS)

Relatamos e discutimos aqui alguns resultados de uma pesquisa que reúne o projeto TEXTQUIM e a Área de Educação Química (AEQ) da UFRGS em torno da linguagem química, em suas diferentes instâncias, em língua portuguesa, incluindo problemáticas de ensino de ciências e da tradução de textos de Química a partir do inglês. São observados terminologias, elementos de textualidade e de enunciação um corpus composto por manuais acadêmicos, artigos de periódicos especializados e textos de revistas de popularização de temas de ciências. Nesse corpus, com auxílio informatizado, numa determinada etapa da pesquisa, buscamos identificar a incidência de palavras e expressões conceitualmente relevantes para o entendimento de determinados

temas da ciência. Partindo do segmento do corpus composto por materiais didáticos utilizados no ensino superior, este trabalho discute problemas e vantagens de uma observação de linguagens especializadas guiada por fatores como frequência de repetição de palavras e relevância conceitual, incluindo elementos ainda pouco alcançáveis por ferramentas de processamento de linguagem, tais como anáfora e retomada lexical por expressões equivalentes ou aproximadas.

USO DE RECURSOS DE INFORMÁTICA EM PESQUISAS NEOLÓGICAS E TERMINOLÓGICAS

Bruno Oliveira Maroneze (USP), Ieda Maria Alves (USP), Mariângela de Araujo (USP)

O Projeto Observatório de Neologismos Científicos e Técnicos do Português Contemporâneo, Projeto Integrado de Pesquisa junto ao CNPq n. 550520/2002-3, criado em 1988, com apoio do CNPq, procurou sempre atender a dois objetivos: por um lado, o Projeto tem buscado observar, analisar e difundir aspectos da criatividade lexical, tanto de unidades lexicais da língua geral como de unidades lexicais especializadas; por outro lado, procura atender às necessidades do desenvolvimento tecnológico e científico por meio da elaboração de trabalhos terminológicos em algumas áreas. Se esses objetivos inicialmente propostos continuam sendo contemplados, os métodos para atingi-los, no entanto, foram sendo adaptados aos recursos que a Informática vem-nos proporcionando. Desse modo, as fichas terminológicas e lexicais, no início sob forma de papel, foram sendo armazenadas em bases de dados (Dbase e, atualmente, Access). Dicionários eletrônicos passaram a auxiliar a determinação do caráter neológico ou não-neológico de uma unidade lexical; textos informatizados possibilitam o uso de programas de busca de unidades lexicais; a disponibilização de textos na Internet tem tornado possível o uso da rede como um grande corpus virtual. Se é incontestável que a Informática possibilita a utilização de recursos cada vez mais eficientes para o estudo do léxico, deve-se também assinalar algumas limitações que o uso das novas tecnologias pode apresentar para esse estudo. Salientamos, especialmente, no que se refere à coleta de unidades lexicais neológicas, o reconhecimento de neologismos semânticos, quer sob a forma de unidades lexicais simples ou de sintagmas, assim como o reconhecimento de composições ou mesmo derivações prefixais grafadas sem hífen (situação problema, não reconhecimento etc.); e, no que concerne à constituição de um corpus representativo, sobretudo nas áreas de especialidade, a busca de textos em formato eletrônico que obedeçam a critérios de qualidade.