

Dicionários utilizáveis no processamento de línguas Legibilidade e organização das informações

Éric Laporte
Universidade Federal do Espírito Santo
Université Paris-Est

O que faz com que um dicionário seja utilizável no processamento de línguas naturais (PLN)? Examinaremos dois requerimentos: a legibilidade das informações e a organização geral, ilustrando o raciocínio com exemplos reais tirados de projetos de dicionários morfossintáticos ou sintáticos.

As principais operações realizadas sobre um dicionário para o PLN são, além do acesso às informações por computador, operações humanas: construção, atualização, verificação, correção.

A legibilidade dos dados fornecidos pelo dicionário é importante porque dela depende a eficácia do trabalho de verificação e modificação das informações por linguistas. Por exemplo, fórmulas com vários níveis de parênteses são um obstáculo à leitura e um verdadeiro convite ao erro.

No que diz respeito à organização geral das informações, estamos a favor do agrupamento de todos os dados relacionados a uma entrada lexical, ou a um conjunto homogêneo de entradas lexicais. Analisamos as dificuldades de manutenção de dicionários que não são organizados em torno da noção de entrada lexical.