

A Análise Morfológica feita por computador necessita de um dicionário de palavras. Convencionou-se que as palavras, no Projeto Lexis, podem ser classificadas como formas e como bases. Uma palavra é uma forma quando ela é armazenada por inteiro e é uma base quando é formada pela junção de duas partes: uma invariável (base) e outra que se flexiona (terminação). Esse dicionário foi implementado usando a estrutura de uma árvore Trie. Trie é um tipo especial de árvore cuja idéia básica consiste em comparar partes da chave de entrada (palavra). A regra geral desse tipo de árvore é que cada letra da palavra determina um nodo da árvore sucessivamente. Foi desenvolvido um protótipo, em linguagem C no ambiente PC-DOS, que realiza a Análise Morfológica de frases da língua portuguesa. O protótipo faz uma análise palavra a palavra da frase identificando as características morfológicas dessas (categoria léxica, gênero, número, grau, etc). Ele mostra todas as categorias léxicas possíveis a que cada palavra pode pertencer. Atualmente, o projeto Lexis encontra-se na fase de migração desse protótipo para a estação de trabalho SUN.(CNPq)