

Oele Koornwinder

Morfologische aspecten van het ideale woordenboek

Een theoretische en empirische studie naar de lexicale samenhang van het Nederlands ten behoeve van een morfologische kennisbank

De centrale stelling van deze studie is dat morfologische structuur een cognitief hulpmiddel is om maximale samenhang aan te brengen in een lexicaal netwerk van kenniseenheden. Hiertoe wordt een analysemethode geïntroduceerd die zich laat omschrijven als Lexicale Kennisrepresentatie door Inductieve Naamgeving (L-KRING). Deze L-KRING-theorie is gemotiveerd door een reeks criteria te geven voor een ideaal model van het lexicon en na te gaan in hoeverre de bestaande lexicontheorieën aan deze criteria voldoen. Vervolgens wordt uiteengezet hoe de beschikbare kennis over de Nederlandse woordbouw integraal kan worden verantwoord met behulp van een op paradigmatische distributiepatronen gebaseerde classificatiemethode. De op deze wijze verkregen inzichten vormen de basis voor een formele beschrijving van de L-KRING-theorie. Deze theorie wordt empirisch onderbouwd door de resultaten te behandelen van een grootschalig lexicografisch onderzoek naar de morfologische structuur van het Nederlands. Voor dit doel is een Morfologische Gegevensbank van het Nederlands opgezet (MGBN). Hierbij zijn alle woorden uit de Grote Van Dale langs semi-automatische weg van morfologische structuur voorzien.

Uit de computationele analyse van deze gegevensbank blijkt dat de Nederlandse woordenschat een sterke samenhang vertoont: de 250.000 woorden uit de Grote Van Dale zijn namelijk uit 80.000 verschillende basislexemen opgebouwd, die zelf weer uit ca. 20.000 wortels, 300 prefixen en 700 suffixen bestaan. De MGBN biedt voor al deze morfemen gedetailleerde informatie over hun combinatiemogelijkheden. Het is daarom een interessante hulpbron voor taalkundig onderzoek en taaltechnologische toepassingen.

key-words: lexicology, lexicography, computational morphology, network model, mental lexicon, automatic knowledge acquisition, knowledge representation system

Oele Koornwinder

Morfologische aspecten van het ideale woordenboek

Een theoretische en empirische studie naar de lexicale samenhang van het Nederlands ten behoeve van een morfologische kennisbank

