

### Analyse structurelle automatique de groupes nominaux anglais

La connaissance précise des relations syntaxiques entre les éléments d'unités linguistiques doit être à la base de toute traduction automatique. Les unités linguistiques de l'anglais sont très souvent représentées par des groupes nominaux. Ceux-ci sont caractérisés par leur élément obligatoire (noyau) qui doit être un substantif. Le noyau peut être accompagné d'un élément facultatif (satellite).

Nous proposons ici un modèle pour l'analyse des groupes nominaux anglais. Dans ce modèle nous travaillons avec cinq niveaux de structure et deux relations binaires, l'extension et la coordination. Selon les niveaux de structure sur lesquels les relations valent, nous distinguons quatre classes de relations de l'extension et cinq classes de relations de coordination. Les deux types de relations sont sub-classifiés.

Le programme d'ordinateur pour l'analyse automatique des groupes nominaux anglais que nous avons développé est basé sur le modèle ci-dessus. Il comprend un programme principal et 13 programmes subordonnés. Destinés pour un ordinateur IBM7094, les programmes sont écrits dans la langue FORTRAN II. Dans l'état où il est, le programme peut analyser des groupes nominaux dans lesquels les satellites précèdent le noyau. Des 906 groupes nominaux qui ont été analysés à titre d'expérience l'ordinateur nous a livré 87,32% d'analyses correctes. L'analyse d'un groupe nominal demandait 0,00424 min.

Les résultats de l'analyse automatique nous servent de point de départ pour des études plus poussées. Nous demandons:

1. Quel rôle jouent-ils dans la traduction automatique?
2. Qu'est-ce qu'ils nous disent sur la réalisation des structures prévues dans notre modèle?

Deux résultats paraissent être importants:

1. L'ordinateur découvre des homonymies structurelles dont le lecteur humain ne s'aperçoit pas;
2. La longueur et la complexité d'un groupe nominal dépend de l'unité linguistique dont ce groupe est un élément.

### Automatische Strukturanalyse englischer nominaler Gruppen

Kenntnis der syntaktischen Relationen innerhalb linguistischer Einheiten einer Sprache ist eine der wichtigsten Voraussetzungen für die automatische Sprachübersetzung. Linguistische Einheiten des Englischen werden sehr oft durch nominale Gruppen manifestiert. Sie enthalten als obligatorisches Element (Nukleus) ein Substantiv, das durch fakultative Elemente (Satelliten) ergänzt werden kann.

Wir schlagen ein Modell für die Analyse dieses Strukturtyps vor, das 5 Strukturebenen ansetzt und mit den binären Relationen Erweiterung und Koordination arbeitet. Mit Hilfe des Konzepts Strukturebene unterscheiden wir 4 Typen von Erweiterungsrelationen und 5 Typen von Koordinationsrelationen. Beide Relationstypen sind unterklassifiziert.

Dieses Modell ist die Grundlage für ein Maschinenprogramm zur automatischen Analyse englischer nominaler Gruppen. Es besteht aus einem Hauptprogramm und 13 Unterprogrammen. Es ist in der problemorientierten Sprache FORTRAN II geschrieben und bewältigt in seinem jetzigen Zustand die Analyse nominaler Gruppen, die nach links erweitert sind. Die Testläufe mit 906 nominalen Gruppen lieferten 87,32% richtige Analysen bei einer Analysezeit von 0,00424 min. pro nominale Gruppe.

Die Ergebnisse der automatischen Analyse werden verwendet zu Untersuchungen über

1. ihre Brauchbarkeit für die automatische Sprachübersetzung,
2. die Realisierungsquote der im Modell enthaltenen Strukturmöglichkeiten.

Folgende Punkte sollen diskutiert werden:

- a. Entdeckung struktureller Homonymien durch die Maschine,
- b. Abhängigkeit von Länge und Komplexität nominaler Gruppen von übergeordneten linguistischen Einheiten

### Automatic Structural Analysis of English Nominal Groups

Comprehensive knowledge of the syntactical relations between the elements of linguistic units is one of the most important prerequisites for automatic language translation. Linguistic units in English are frequently represented by nominal groups. They are characterized by their compulsory element (nucleus) which must be a noun; it can be expanded by optional elements (satellites).

In our paper we propose a model for the structural analysis of English nominal groups. In this model we distinguish five structural levels and two binary relations, expansion and coordination. According to the structural level on which they hold, we postulate four types of expansion relations and five types of coordination relations. Both types of relations are sub-classified.

A machine program for automatic analysis of English nominal groups is based on this model. It consists of a main program and 13 sub-programs. The whole complex of programs is written for a computer, model IBm7094. The programming language used is FORTRAN II. In its present stage it analyses nominal groups in which the nucleus is preceded by its satellites. Tests have been made with 906 nominal groups. They yielded 87,32% of correct analyses. The analysis of one nominal group took 0,00424 min.

The automatic analysis results have been further examined. We asked two questions:

1. How important are they for automatic language translation?
2. What do they tell us about the adequacy of our model?

Our attention has been focused on two facts:

1. Automatic language analysis discovers structural homonymies which the human observer doesn't become aware of;
2. Length and complexity of nominal groups depend on the linguistic units in which they occur.