GENERATING SEMANTIC STRUCTURES IN EUROTRA-D
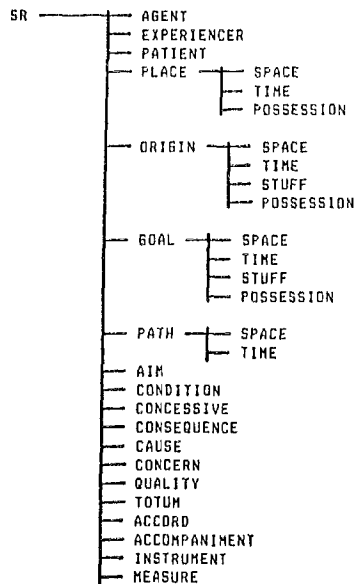
ERICH STEINER
IAI - EUROTRA - D
MARTIN-LUTHER-STRASSE 14
D - 6600 Saarbrücken
WEST GERMANY

1. The following paper is based on work done in the multi-lingual MT-project EUROTRA, an MT-project of the European Community.

Analysis and generation of clauses within the EUROTRA-framework proceeds through the levels of (at least) EUROTRA constituent structure (ECS), EUROTRA relational structure (ERS) and interface structure (IS) (cf. Arnold/des Tombe/Jaspaert.1985 and the EUROTRA-REFERENCE MANUAL, version 1, revision 0).

At IS, labelling of nodes consists of labellings for time, modality, semantic features, semantic relations and others. In this paper, we shall be concerned exclusively with semantic relations (SRs), to which we shall also refer as "participant roles" (PR). A list of EUROTRA SRs as given in ELS-3 is reproduced below:

```
SR ————┬— AGENT
        ├— EXPERIENCER
        ├— PATIENT
        ├— PLACE ———┬— SPACE
        │           ├— TIME
        │           └— POSSESSION
        │
        ├— ORIGIN ——┬— SPACE
        │           ├— TIME
        │           ├— STUFF
        │           └— POSSESSION
        │
        ├— GOAL ————┬— SPACE
        │           ├— TIME
        │           ├— STUFF
        │           └— POSSESSION
        │
        ├— PATH ————┬— SPACE
        │           └— TIME
        ├— AIM
        ├— CONDITION
        ├— CONCESSIVE
        ├— CONSEQUENCE
        ├— CAUSE
        ├— CONCERN
        ├— QUALITY
        ├— TOTUM
        ├— ACCORD
        ├— ACCOMPANIMENT
        ├— INSTRUMENT
        └— MEASURE
```

According to current EUROTRA legislation, these SRs are assigned to dictionary entries of verbs (and other word classes, which will be disregarded in this paper) by coders, and through these entries to clauses in a pattern matching process.

This approach, while certainly valid in principle, leads to the problem of inter-coder-consistency, at least as long as the means for identifying SRs are paraphrase tests for SRs. In EUROTRA-D, we have for some time now been experimenting with a set of SRs, or PRs, which are identified with the help of syntactic criteria. This approach will be outlined in the following. Its roots are to be found in the literature on semantic cases in general (cf. Fillmore.1968 and 1977, Starosta 1977, Somers.1983) and more particularly, in ideas from Systemic Linguists (cf. Fawcett.1980 and forthcoming, Halliday.1967-68 and 1985, Steiner.1983 and 1985, Winograd. 1983).

2. Definition of basic concepts

To start with, we need a small set of basic concepts, which will be defined in the following. There is a basic distinction between Participant Roles and Circumstantial Roles on the semantic level, which the ELS-3 specifications do not contain. There is the further distinction between inherent and non-inherent roles, which is also not explicitly made in ELS-3 and may be necessary at some point.

Participant Role: A participant role is a semantic constituent of a clause which is syntactically realized as a complement rather than as a modifier.

Circumstantial Role: A circumstantial role is a semantic constituent which is, on clause level, realized as a modifier rather than as a complement

Obligatory Participant Role: A semantic constituent which is, on clause level, realized as obligatory complement

Optional Participant Role: A semantic constituent which is, on clause level, realized as an optional complement

Inherent Role: Inherent roles are obligatory participant roles and those participant roles which, if they are not realized in a clause, lead to look up in the preceding text or situation for a referent. Cf.: (20) David was watching.

Non Inherent Role: Non inherent roles are those participant roles which, if they are not realized in the clause, do not lead to look up. Cf.: (21) Devid is eating

3. Definitions of semantic relations and types of process

The first essential step is to define certain key PRs as far as possible through syntactic criteria. Some of these definitions are given in the following:

SR/PR: Phenomenon
Syntactic realization: - dependent clause in the function of syntactic "direct object", introduced by "that" or "if";

- dependent ing-construction in the syntactic function of "direct object".

SR/PR: Location
Syntactic realization: - complement of place, origin or direction (realized by Adverbial Group, Prepositional Group or embedded clause)
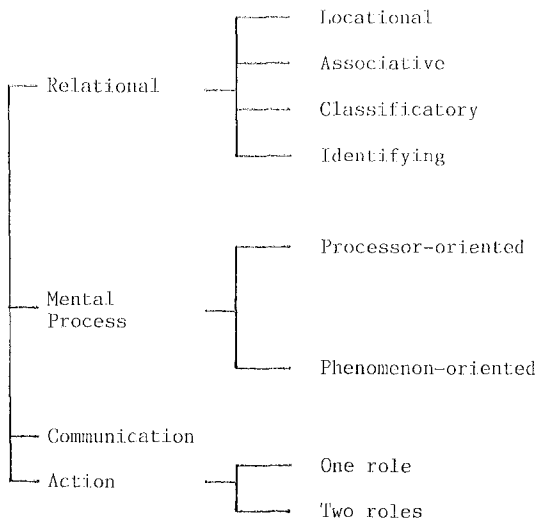
SR/PR: Identifier
Syntactic realization: - NG in function of subject complement
- ADJ-group with the adjective in the superlative
- embedded clause in function of subject complement

These are just a few characteristic examples. The overall number of PRs is around 15. All these PRs have, apart from syntactic criteria for their identification, notional definitions and paraphrase tests. These, however, are only used in the cases, where syntactic criteria do not lead to an unambiguous assignment.
An identified sub-set of the PRs is used to identify "types and process". These, in turn, serve as entry conditions into very limited sub-sets from the overall set of PRs, so that the choice of PRs for the coder is always limited to a small number of between 2 and around 6 cases from the overall set. The types of process are given in the following:

Types of Process

- Relational
  - Locational
  - Associative
  - Classificatory
  - Identifying
- Mental Process
  - Processor-oriented
  - Phenomenon-oriented
- Communication
- Action
  - One role
  - Two roles

4. A procedure for the assignment of semantic relations

The following procedure is used for the assignment of semantic relations by dictionary coders.

FLOWCHART : How to assign Semantic Relations
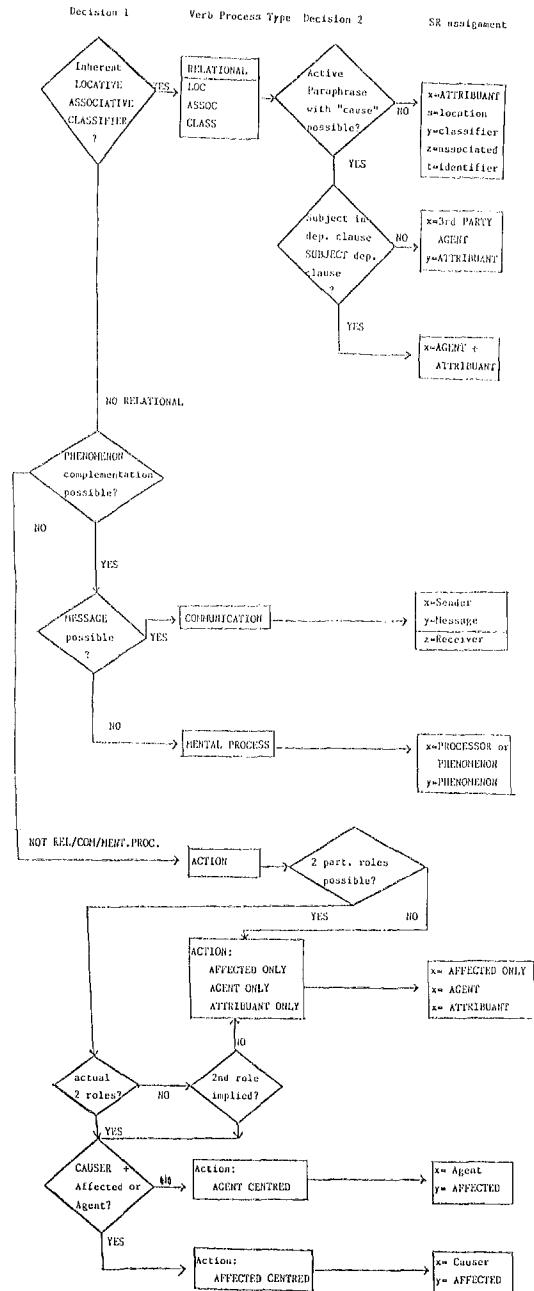
5. Conclusion

The statements in this paper are necessarily very sketchy and brief. More detailed information will be given in the oral presentation.
The assignment of semantic relations, as presented here, has been worked out for German in some detail. The available literature for English, especially in the work of Fawcett and Halliday, shows that the treatment of English in the same way does not present a major problem. It seems to be reasonable to assume that the method outlined constitutes a worthwhile way to explore for a multi-lingual MT-project such as EUROTRA.

6. Bibliography

1. Arnold/des Tombe/Jaspaert. 1985. ELS-3. EUROTRA Linguistic Specifications. Version 3.

2. Blatt/Luckhard/Truar. 1984. Evaluation of LS-1 legislation concerning semantic relations (ETL-4-D)

3. Fawcett, R.P. 1980. Cognitive linguistics and social interaction.
   Heidelberg: Julius Groos

4. Fawcett, R.P. forthcoming. The semantic of clause and verb for relational processes in English. in: Fawcett/Halliday. eds. forthcoming. New developments in Systemic Linguistics.
   London : Frances Pinter

5. Fillmore, C.F. 1977. The case for case. in: Bach/Harms. eds. 1968. Universals in linguistic theory. New York: Holt, Rinehart & Winston. 1 - 88

6. Fillmore, C.F. 1977. The case for case reopened. in: Cole/Saddock. eds. 1977. Syntax and Semantics, Vol. 8. New York: Academic Press. 59 - 81

7. Halliday. M.A.K. 1967-68. Notes on transitivity and theme in English. in: Journal of Linguistics, 3.1. 1967, 37-81; 3.2. 1967, 199-244; 4.2. 1968, 179-215

8. Halliday. M.A:K. 1985. An introduction to functional grammar.
   London : Edward Arnold

9. Ruus/Spang-Hansen. 1982. Argument relations and predicate types for EUROTRA. (ET-10-DK)

10. Somers. H.L. 1983. An investigation into the application of the linguistic theories of valency and case to the automated processing of natural language. UMIST PhD-thesis

11. Starosta, St. The one/sent solution. Trier: LAUT

12. Steiner, E. Die Entwicklung des Britischen Kontextualismus.
    Heidelberg : Groos

13. Steiner, E. 1985. Working with transitivity: system networks in semantico-grammatical descriptions. in: Benson/Greaves. eds. 1985. Systemic perspectives on discourse Vol. 1
    Norwood, N.J. : Ablex

14. Winograd, T. 1983. Syntax as a cognitive process. Vol. 1.
    New York : Willey