

## Structuring of information

In the most important theory on coherence, the Rhetorical Structure Theory (Mann & Thompson 1988), and in discussions about this theory (for example Hovy 1990, Sanders, Spooren & Noordman 1992, Moore & Pollack 1992 and Sanders & Spooren 1999) several issues are left open. In various sets of coherence relations a distinction is drawn between relations that refer to the world (for example cause-consequence) and relations that express the writer's intention (for example motivation). However, it remains unclear to what extent relations of the first type can also be of an intentional nature.

Moreover, it seems difficult to classify certain relations under one of both types, for example the elaboration as in: "Do you see that man over there? That's my father." A theory of coherence relations should not only account for these diverse types of relations, it should also explain why relations can be interpreted so differently. In a reader survey (Renkema 1989) for example, the relation between the clauses in "Pete is the manager. Therefore you should discuss the matter with him" was labeled as volitional cause, motivation, justification and conclusion.

A third problem which still lacks an appropriate solution concerns the occurrence of relation markers. In the last example the connective can be omitted. Nonetheless, a certain cohesion strategy then seems to be preferred, in this case a 'last-out-first-in' principle: "Pete is the manager. With him you should discuss the matter."

This project aims to contribute to research into the types of coherence relations, the ambiguity of those relations and the factors that determine 'markedness'. The project's contribution to the framing of a theory is threefold:

1. Picking up on German text theories, Harweg's substitution approach (1968) and Weinrich's score approach (1976), an attempt will be made to design a model of coherence relations from a three-way perspective: syntagmatic-semantic-pragmatic. The Multi-Level Hypothesis then will be extended with a component accounting for coherence relations that are described as complement relations in more traditional analyses of complex sentences. It is expected that this model can more justly lay claim to descriptive adequacy than existing models.
2. Earlier reader research into labeling of coherence relations showed that ambiguity could partly be explained by supposing a semantic and a pragmatic reading. This research will be continued. As a follow-up, it will be investigated to what extent ambiguity can be answered for in a theory that takes readers' expectations into account as a result of text type and reading strategies.
3. Assuming the existence of an interaction between the 'markedness' of rhetorical relations and (referential) cohesion strategies, the earlier developed descriptive model for referential cohesion (Renkema 1996, 1998) will be used to analyze rhetorical relations (whether or not marked through the use of connectives) in context.

The project's procedure is as follows. In the preparatory phase an inventory of all coherence relations that are described in the literature will be taken, including complement relations. Then a corpus of short texts (1,000 words maximum) will be built up, which makes it possible to analyze relations in full context. Marked and

unmarked relations will be included, and within the marked relations a distinction will be drawn between connectives and relation determining NPs, such as ‘This is the reason that’.

The entire corpus is considered to contain approximately 1,000 texts. Four subcorpora will be built up consisting of informative texts (teletext pages, short news reports), narrative texts (from well-known short story volumes), instructive texts (passages from manuals and explanations accompanying forms) and argumentative texts (letters to the editor, opinion page texts).

In the interaction between stocktaking of literature and corpus research, a three-way model (including syntagmatic, semantic and pragmatic relations) will be developed which, on the basis of new input, will be refined and tested for applicability and coding reliability. Subsequently, in experimental research readers will be asked to name coherence relations in full context. The tasks will include free coding, choosing from relations of one type (for example of the semantic type volitional cause, non-volitional cause, result/consequence) and choosing from relations of several types. Subjects will also be asked to account for the relation they chose.

Depending on the results, the focus will be on possible explanatory factors such as nature of ‘markedness’ or text type. Simultaneously the marked and unmarked relations will be described using the descriptive model for cohesion. It is to be expected that for this purpose the model have to be refined and tested for coding reliability. The refined model will then be used to determine if and if so which type of coherence exists between marking and certain cohesion strategies.

It is proposed to do parts of this research in consultation with the designers of the Rhetorical Structure Theory (William Mann, Dallas, and Sandra Thompson, Santa Barbara). The societal importance of this theory developing research can be indicated as follows. In writing education notions as ‘deficient coherence’, ‘sudden jump of thought’ or ‘illogical structure’ are dealt with in a rather impressionistic or ad hoc manner. Illustrative is the writing advice to use connectives if that is necessary to improve clarity (without giving criteria for ‘necessity’ or ‘clarity’). The proposed research into the levels of coherence and the way in which readers construct coherence may contribute to the ‘structuring’ component in writing education.