Rhetorical structure and thematic structure in text generation

Manfred Stede
KIT Project Group
Technische Universität Berlin / Germany
stede@cs.tu-berlin.de

1 Rhetorical trees in text generation

In text generation, rhetorical trees are a popular intermediate structure used in between text planning and surface-linguistic realisation. In particular, Rhetorical Structure Theory (RST, [Mann, Thompson 1988]) trees have become very influential, mostly because the relations are defined in terms of intentions/beliefs and thus invite being formalized and used in text planning (e.g., [Hovy 1988], [Moore, Paris 1993]), which is commonly regarded as the first step of the generation process.

On the other hand, the second step — linearizing RST trees into a sequence of sentence plans — has received much less attention. Amongst the exceptions, Scott and deSouza [1990] suggested a set of heuristics for signalling rhetorical relations unambiguously on the surface. Knott [1991] described a generator that incorporates some of these heuristics but is also critical of some others. He proposed a bottom-up procedure that employs a set of patterns associated with discourse markers (connectives signalling the relation) in order to determine an overall consistent breakdown of the rhetorical tree into clause-size chunks, linked by sentence boundaries and/or discourse markers. Rösner and Stede [1992] suggested a similar algorithm. The two major problems with such straightforward procedures are that they rely upon a somewhat ill-defined notion of "size" of embedded material (subtrees), and that the contents of the propositions are rather neglected — as we will see, quite often contents need to be considered to determine an appropriate thematic structure.

Linearizing an RST tree involves, among other tasks, determining linear order on all linguistic levels: that between clauses (roughly, between leaf nodes of the rhetorical tree) and that within clauses. Unless, of course, the RST tree is seen as already determining the order in which the minimal units are to be presented; this is the case for instance in the systems of Hovy [1988], Knott [1991], Marcu [1997]. Other text planners, however, advocate making a clear distinction between content selection and content ordering (e.g., Gossip [Iordanskaja 1992]). This indeed seems a necessary step; Marcu [1997, p. 248], for example, concludes that a fixed-order RST tree constitutes a simplification and that factors such as focus, distribution of given and new information, and high-level pragmatic and intentional constraints should be taken into account. In short, the move from schema-based generation to rhetorical trees has opened up a range of ordering tasks that earlier went unnoticed. In "pre-RST times", ordering could quite conveniently be handled by the schemata; McKeown [1985], for instance, defined schemata that partially ordered the sequence of propositions, and used an algorithm that computes permissible focus shifts in order to decide on locally remaining alternatives.

2 Thematic structure

Linear order in a text is, of course, determined by many factors. One of them, to which I wish to draw attention here, is a felicitous 'thematic structure'. To illustrate its role, consider the two text versions in Figure 1 (taken from Fries [1981]), the second of which is less acceptable due to suboptimal theme choices, as Fries argues. Themes are underlined; Fries' definition of this notion will be given in Section 4. Notice the lack of "flow" in the second version, for instance in lines 7 and 8, where reversing CAUSE and CONSEQUENCE, and moving the phrase *in World War II* to the end of the sentence has caused disruptions. I have undertaken an RST analysis of the original text, which is shown in Figure 2. In a nutshell, the analysis posits that the "point" of the paragraph is stated at the very beginning: The nature of U.S. participation in the two world wars was different. Then, three items of evidence are given: number of theaters, role of the

Navy, and role of airpower.

Switching now the perspective from analysis to generation, assume that a text planner has produced the tree in Figure 2: relations and representations of propositions, no reference to a particular linear order (i.e., the numbers used for the propositions in the figure are not significant). In producing this tree, the planner has already ruled out other options for presenting the material, for instance an organisation that first gives a block of information about US participation in WW I, followed by a block of information about WW II. Now, given the tree in Figure II, the remaining text generation problem is to produce a text that — as far as thematic structure is concerned — is close to the first version in Figure 1, and that avoids the mistakes made in the second version.

In generation, only few systems have made reference to thematic structure (e.g., Gossip [Iordanskaja 1992], KOMET [Bateman et al. 1998]). This is surprising at first sight, but it can be explained at least in part by the choice of domain in most RST-based generators of the last decade: Instructional or other task-oriented discourse was dominant in the applications (for a number of good reasons), e.g. in the systems of [Rösner, Stede 1992], [Paris et al. 1995], [Kosseim, Lapalme 1995], [Vander Linden, Martin 1995]. Here, the domain or task structure to a large extent determines linear order in the text; loosely speaking, the RST trees are typically populated with SEQUENCE relations very close to the root node. Also, the individual steps in an instruction are usually self-contained and not rhetorically complex. Thus, not too much can go wrong when linearizing an RST tree representing instructions.

When leaving the linearisation-friendly domains, questions of thematic structure become acute, though. More specifically, the following three questions arise, to which I will attend in turn.

1. How much linear order is already coded implicitly in the rhetorical tree?
2. How can the remaining clause ordering be performed?
3. How can the clause-internal ordering be determined?

3 Implicit ordering constraints in rhetorical trees

Mann and Thompson [1988], in considering RST relations as "pre-realisational", leave issues of signalling them at the surface aside, and regarding ordering, they merely give a list of typical, frequently observed orderings. While such tendencies can be used as rough heuristics for linearisation (e.g., [Moore and Paris 1993]), they are clearly too static when we turn to the task of constructing a successful thematic development. Assuming the other extreme, however, that a rhetorical tree leaves linear order totally unconstrained, seems not right, either.

First of all, we adopt the general assumption that nuclei and satellites of a relation are to be realized in adjacency. That is, the tree in Figure 2 cannot be linearized with, for instance, proposition [7] in between propositions [8] and [9]. For the subtree [7,8,9], the search space of "structurally permissible" orderings is therefore {[7-8-9], [7-9-8], [8-9-7], [9-8-7]}. At the same time, the constraint of nuclei/satellite adjacency leaves many choices for moving blocks of text around; for example, it does not preclude realizing the subtree [7,8,9] after [10,11,12,13].

A second source of ordering constraints stems — potentially — from the text planner. If the text is argumentative in nature, then linear order is likely to be closely related to its argumentative structure and thus largely to be decided by the planner. In such cases, order can to some extent be fixed on the global level but also locally; for instance, a CONCESSION is often rhetorically more emphatic if the satellite (the conceded element) is presented before the nucleus (the point made "anyway"). Similarly, relizing a CONTRAST (which RST views simply as a multi-nuclear relation, implying no specific ordering), in English prototypically done with the conjunction *but*, often displays strong focusing effects [Chafe 1976; Umbach and Stede in prep.], and thus the order of the conjuncts is by no means arbitrary. For illustration, consider this example discussed by Elhadad and McKeown [1990], where the 'argumentative intent' is said to be responsible for the ordering, which must therefore be specified in the rhetorical tree:

He is smart, but he failed the exam. Let's fire him.
He failed the exam, but he is smart. Let's hire him.

Finally, apart from such intention-driven orderings, several rhetorical relations quite clearly imply, by definition, how the spans are to be ordered. Considering the relations given by Mann and Thompson, I would single out BACKGROUND, INTERPRETATION, EVALUATION, SUMMARY, and possibly RESTATEMENT, SOLUTIONHOOD and OTHERWISE. The SEQUENCE relation displays "strong prefer-

1. Although the United States participated heavily in World War I (1), the nature of that participation was fundamentally different from what it became in World War II (2).
2. a The earlier conflict was a one-ocean war for the Navy (3) and a one-theater war for the Army (4);
3. b the latter was a two-ocean war for the Navy (5) and one of five major theaters for the Army (6).
4. c In both wars a vital responsibility of the Navy was escort-of-convoy and anti-submarine work (7),
5. a but in the 1917–1918 conflict it never clashed with the enemy on the surface (8);
6. b whilst between 1941 and 1945 it fought some twenty major and countless minor engagements with the Japanese Navy (9).
7. a American soldiers who engaged in World War I were taken overseas in transports (10) and landed on docks or in protected harbors (11);
8. b in World War II the art of amphibious warfare had to be revived and developed (12), since assault troops were forced to fight their way ashore (13).
9. a Airpower, in the earlier conflict, was still inchoate and almost negligible (14);
10. b in the latter it was a determining factor (15).

1. Although the United States participated heavily in World War I (1), the nature of that participation was fundamentally different from what it became in World War II (2).
2. For the Navy the earlier conflict was a one-ocean war (3) and for the Army a one-theater war (4);
3. the latter was a two-ocean war for the Navy (5) and one of five major theaters for the Army (6).
4. A vital responsibility of the Navy was escort-of-convoy and anti-submarine work in both wars (7),
5. but it never clashed with the enemy on the surface in the 1917–1918 conflict (8);
6. whilst some twenty major and countless minor engagements were fought with the Japanese Navy between 1941 and 1945 (9).
7. American soldiers who engaged in World War I were taken overseas in transports (10) and landed on docks or in protected harbors (11);
8. since assault troops were forced to fight their way ashore (13), the art of amphibious warfare had to be revived and developed in World War II (12).
9. Airpower, in the earlier conflict, was still inchoate and almost negligible (14);
10. it was a determining factor in the latter (15).

Figure 1: Two versions of sample text (from Fries [1981])

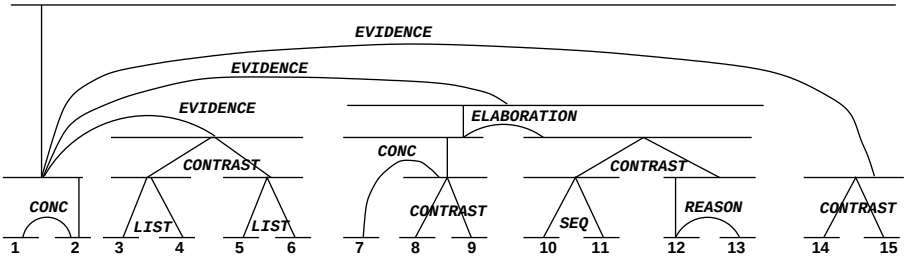

Figure 2: RST analysis of text in Figure 1

ences", especially if joining more than two nuclei, but is less strict; the situation is similar for ELABORATION. The remaining relations, however, essentially allow a free order: ANTITHESIS, CONTRAST, CONDITION, CONCESSION, CIRCUMSTANCE, ENABLEMENT, EVIDENCE, JUSTIFY, MOTIVATION, PURPOSE, and the CAUSE cluster.

Generally problematic for ordering are multi-satellite relations, which occur quite often when analyzing "everyday text" but are typically excluded from implementations, and instead re-structured as binary trees, which makes processing much easier in several respects. The text in Figure 1 demonstrates that this is a simplification: The EVIDENCE relations all share the same nucleus, and an artificial binary analysis would not do justice to the text. In which order the three satellites are to be presented, however, can not be easily read off the tree.

4 Ordering text units

After taking the above-mentioned three types of constraints into account, many decisions for order-

ing the units of the text are still to be made. The problem is aggravated by the fact that in RST–based generation (as opposed to the original proposal by Mann and Thompson [1988]), relations are commonly assumed to also hold on the sub-clausal level (e.g., [Knott 1991], [Rösner, Stede 1992], [Vander Linden, Martin 1995]) and can link, for instance, a prepositional group to a clause: *For best results, use brand new oil.* Thus, there are even more ordering decisions among 'minimal units' to be made.

One approach taken in response to the trees' under-specification w.r.t. ordering was Hovy and McCoy's [1989] suggestion to build up 'focus trees' in tandem with rhetorical trees. For every relation node in the rhetorical tree, the corresponding node in the focus tree would list the daughter nodes representing propositions that the focus of attention can move to. The focus-shift rules build upon those used by McKeown [1985]. This is a promising approach in particular when choosing how to order a set of ELABORATIONS or CIRCUMSTANCES, where the goal is to avoid going back and forth between the same topics in unnecessarily complicated orders.

A second approach is to seek general ordering principles independent from focus shifts. Iordanskaja [1992] suggested 11 such rules for ordering adjacent spans. When viewing her cases from an RST perspective, it turns out that the majority would be handled by the RST relation ELABORATION; they are specific instances of what it means to elaborate. The importance of developing more fine-grained accounts of ELABORATION (and also CIRCUMSTANCE) has been stressed also by Knott [1991], and this is particularly relevant for issues of linear order. The three non-elaborating cases of Iordanskaja are: (1) Crossing of relations is not allowed, except for certain comparisons: *Two users made use of the system. Martin edited some files. Jessie ran compilers. The files edited by Martin are non-sensitive.* — This example seems indeed problematic for an account faithful to the RST principles. (2) Particularly important information must be placed at the beginning or end of the text. — This would properly be decided by the text planner. (3) Principle of symmetry: When describing attributes of objects, introduce the attributes in the same order as the objects. — This principle is at work in our sample text in Figure 1, where the contrast between the first and the second war is stated several times, and it would be awkward if the order of the two were changed at some point.

A different ordering principle is at work at the very beginning of the text. In general, the construction *Although A, B* can usually be replaced by *B, although A*. For example: *Although it was raining, we went for a walk. — We went for a walk, although it was raining.* With the clauses (1) and (2), though, this replacement is very difficult, if not impossible. A reversed verbalisation would be *The nature of U.S. participation in WWII was fundamentally different from that in WWI, although the U.S. participated heavily in the earlier war.* The sentence sounds awkward because the *although* does not at all introduce a counter-argument to that given in the matrix clause, as it would do in a "standard" CONCESSION (cf. the discussion of 'internal' versus 'external' CONCESSION in [Grote et al. 1997]). Rather, the function of the whole sentence is to direct the reader's attention to the matrix clause, where the main argument of the text is stated, for which the following material will provide evidence. This is a matter of thematic structure: The CONCESSION operates on the textual, not on the semantic level. Notice this is not dependent on the subordination introduced by *although*, as the paraphrase '(1) *but* (2)' is perfectly valid. (And again, the reverse order '(2) *but* (1)' is not.) For situations of this type we require a rule that favours placing the satellite of a CONCESSION (or similar relation) in front of the nucleus, in case the two together form the nucleus of a larger span whose satellites are EVIDENCES (or ELABORATIONS, CIRCUMSTANCES).

A similar situation holds in the subtree [10,11,12,13]. In the second version of the text, the satellite of the REASON relation [13] has been moved in front of the nucleus [12], which results in an odd sequence. Again, we can employ a rule stating that for CONTRASTS with complex nuclei, the embedded (contrasting) nuclei should be realized in adjacency. This would rule out the [10-11-13-12] ordering; notice, however, that the second version can also be saved by thematising the *in WWII* constituent: *. . . and landed in docks or in protected harbors; in WWII, since assault troops were forced to fight their way ashore, the art of amphibious warfare had to be revived and developed.*

In summary, some orderings can be performed in text planning by working out finer grained ELABORATIONS and CIRCUMSTANCES and using focus shift rules, which make use of a knowledge base that knows how the verbalized entities relate to

one another. Other decisions can be made following the goal to produce parallel structures, for example when verbalizing CONTRASTS or LISTS. For the remaining work, however, it is in general not enough to look at individual nucleus–satellite pairs alone. Inspecting the configurations of relations and checking for their possible realisation options is necessary and requires more complex patterns/rules, such as the ones we have mentioned.

5 Clause-internal ordering: choosing themes

Besides fixing the sequence of text units, the second aspect of thematic structure is in choosing themes of individual clauses, which amounts to propagating the overall theme development down to the clause level. (The precise division of labor between ordering units and clause-internal ordering depends of course on the notion of 'minimal unit' one adopts for the RST tree.) Within the clause, various syntactic means can be employed to realize the theme: choose the subject; topicalisation; clefting; choose active or passive voice; choose a verb that allows for choosing the 'right' subject, i.e. one that takes an appropriate perspective on the situation to be verbalised.

Much discussion has been devoted to teasing apart the notions of 'given', 'topic', 'subject', and 'theme' — see for example [Chafe 1976] for an overview. For our purposes here, we sympathize with the systemic-functional tradition and the specific proposal by Fries [1981] who argues that the two dichotomies theme/rheme and given/new should be kept distinct. Fries characterizes 'theme' as "the point of departure of the sentence as message", which in English co-incides with sentence-initial position (whereas in Japanese, for instance, it is often marked with the *wa* particle). Very often the theme is also the given, but not always. In this view, all sentences have themes but not all have given information. The central criterion employed by Fries is that the theme is placed at the beginning of the sentence *as a result of choice*. As a consequence, constituents that are syntactically required to be initial are considered only 'weakly thematic'. The conjunction *but*, for example, must be initial; the closely related *however* need not. Thus, if *however* appears sentence-initial, it is more thematic than a *but*. Notice this is an example of a *textual* theme, which also demonstrates the difference to the given/new distinction.

Sometimes, themes are chosen for reasons of giving *emphasis* to a specific piece of information. But when goal-directed emphasis is absent, there are both good and bad ways of choosing themes in context: Good ways produce a "flowing" text, bad ones do not. An often-cited 'root' of work on theme are the three patterns of thematic development proposed by Daneš [1974]: (1) the theme remains constant for a series of sentences; (2) the rheme of one sentence becomes the theme of the next; (3) all themes in a paragraph are derived from a common, more abstract theme. For the first and third type, we often find evidence where the text topic or genre influences thematic choice for entire paragraphs. Ramm and Villiger [1995], for example, show an excerpt from a travel guide describing a city. In the 8 sentences, all the themes refer to spatial conceptualisations — the text is organized according to what they call a "spatial chaining strategy." In their second example, an argumentative text (which also deals with spatial information) displays object themes in the first half, the counter-argument section, and textual themes in the second, pro-argument half: *moreover, besides, last but not least ...*

This is also one of the hypotheses put forward by Fries [1981]: Choosing a 'type' of theme informs the reader of the type of organisation underlying the text. In the text in Figure 1, the type of organisation is clearly the contrast between earlier war and later war, which therefore is recurrently thematic. A defect resulting from violating this strategy can be noticed in propositions [7,8,9]. In the first version, [7] provides the theme *in both wars*, from which the themes of [8] and [9] are derived (earlier/later war). In the second version, all three clauses thematise the Navy, which does not lend support to the point of CONTRAST made here. The importance of text-global themes is stressed also by Bateman et al. [1998], who argue that the genre of encyclopaedia entries about artists favours a strategy of theme selection that treats the artist as "macro theme" of the text and hence favors its selection for clause themes.

But when no text–global considerations resulting from the specific genre are applicable, theme choice is decisively influenced by the linguistic context — the notion is a dynamic one, as especially Daneš' second pattern indicates. Williams [1990] discusses a sample text (slightly shortened here) that happens to nicely illustrate the power of this pattern:

Astonishing questions have been raised by scientists exploring black holes in space. A black hole is created when a dead star collapses into a point perhaps no larger than a marble. So much matter compressed into so little volume changes the fabric of the space around it in puzzling ways. Most disturbing is the fact that...

Notice that choosing a passive in the first sentence allows for thematising *astonishing questions*, which is not only a "grabber" for the reader but also the 'macro-theme' of the text. The end of the sentence leads to *black holes*, which is the immediately adjacent theme of sentence 2, again due to choosing a passive. This sentence closes with *a point no larger than a marble*, which is taken up in the theme of sentence 3 and moved over to *puzzling ways*, which the last sentence once again takes up with the phrase *most astonishing*.

Unfortunately, though, Daneš' theme development patterns are in no way prescriptive, and there is little evidence as to on what grounds one should choose between the two in a specific situation of text production. Fries [1981] and several other authors noted that texts usually do not exhibit one consistent pattern but often switch between them. The question that is particularly interesting for our topic is whether a connection between theme development and rhetorical structure can be perceived.

Exploring this issue, Matthiessen and Bateman [1991] briefly investigated the link between thematic structure and RST. They suggested that themes are often the "points" of rhetorical relations, which they illustrated with the following CONTRAST example: *In his mind, he.... But in reality, Mehmet....* There are two instruments for signalling the contrast here: the *but*, and the foregrounding of the points of contrast. With similar examples for ELABORATION, Matthiessen and Bateman state that "transitions between partitions motivated rhetorically redistribute thematic statuses accordingly." Our sample text in Figure 1 (i.e., the "good" version) indeed supports this analysis: the points of contrast are consistently thematic. Only, it is not clear how the notion of "point of a relation" would generalize to the other relations, such as the CAUSE cluster: One state of affairs causes another, both can be complex, so what is the "point"?

It thus seems that only few RST relations yield specific instructions on choosing themes. To some extent, this is due to the rather coarse granularity of relations such as ELABORATION, as mentioned earlier. Wanner [1994] stressed the need for more fine-grained coherence relations that allow for systematic links between relations and the lexical material offered by a language — which also involves issues of surface ordering. In order to build a Daneš-type-2 chain of themes and rhemes (where 'rheme' really refers to the end of the clause rather than to 'non-theme'), the connections between the components of the propositions need to be computed. Furthermore, the decisions on theme/rheme need to be recorded in a discourse history, so that theme choice for the next clause can take the prior decision into account, as well as the presence of a 'macro-theme', in case the text or a particular part of it exhibits one.

6 Conclusion: Incremental linearisation

Thematic structure in text is an inherently dynamic affair, specific to clauses in their context. Choosing the beginnings and endings of sentences plays the vital role for creating "flow", as guide books on good writing teach us (e.g., [Williams 1990]). Rhetorical structure, on the other hand, is largely static in nature. It can be determined by top-down processes, whereas thematic structure cannot: It only arises during linearisation, when the options in surface realisation are determined and weighted against one another.

If a text generator is to construct adequate theme development, we require more sophisticated linearisation procedures than those in use today. The picture emerging from our discussion above is one of incremental linearisation, where the rhetorical tree is consumed piecemeal and a history describing relevant features of the text generated so far is kept. The process can be sketched as follows:

1. The text planner produces a rhetorical tree. Some of the relation nodes may already be annotated with directives on ordering the spans, for reasons of argument structure, nature of rhetorical relation, etc. (cf. section 3)

2. The beginning of the text, i.e., the first leaf to be generated, is chosen.

3. By inspecting the neighbouring relations (cf. section 4), the ordering annotations, the contents of the propositions, and the discourse history, one or a few chunks of the tree are consumed and their linear order determined.

4. For each successive chunk, a clause theme is chosen, and a sentence plan is constructed.

5. The plan is passed to the surface realizer, the thematic decisions are recorded in the discourse history.

6. Go to step (3) for consuming more parts of the tree.

Work on fleshing out such a model is under way. A central design decision was to build a dedicated discourse marker lexicon, which assembles the information associated with the various connectives and adverbs available for expressing a rhetorical relation. This lexicon will hold the information on what span orderings are allowed with a specific discourse marker, what syntactic constraints are associated, how much material can be embedded, etc.; it will be employed as a central resource in the sentence planning step (4) above, as explained in [Grote, Stede 1998].

References

J. Bateman, K. Rondhuis. "Coherence relations: towards a general specification." In: *Discourse Processes* 24(1), 1997.

J. Bateman, E. Teich, M. Alexa. "Generic technologies for selective information presentation: an application of computational linguistic methods." In: P. Fankhauser, M. Ockenfeld (eds.): *Integrated Publication and Information Systems: 10 years of research and development.* GMD Forschungszentrum Informationstechnik, Sankt Augustin, 1998.

W. Chafe. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view." In: C. Li (ed.): *Subject and Topic*, pp. 27-55. New York: Academic Press, 1976.

F. Daneš. "Functional sentence perspective and the organization of the text." In: F. Daneš (ed.): *Papers on Functional Sentence Perspective*, pp. 106-128. The Hague: Mouton, 1974.

M. Elhadad, K. McKeown. "Generating Connectives." In *Proc. of the 13th Conference on Computational Linguistics*, Helsinki, 97-101, 1990.

P. Fries. "On the status of theme in English." In: *Forum Linguisticum* 6.1, pp. 1-38, 1981.

B. Grote, N. Lenke, M. Stede. "Ma(r)king concessions in English and German." In: *Discourse Processes* 24(1), 1997.

B. Grote, M. Stede. "Discourse marker choice in sentence planning." In: Proceedings of the Ninth International Workshop on Natural Language Generation, Niagara-on-the-Lake, Canada, 1998.

E. Hovy. "Planning coherent multisentential text." In: Proceedings of the 26th ACL Conference, Buffalo, 1998.

E. Hovy, K. McCoy. "Focusing your RST: a step toward generating coherent, multisentential text." In: Proceedings of the 11th Annual Conference of the Cognitive Science Society, Ann Arbor, 1989.

L. Iordanskaja. "Communicative structure and its use during text generation." *International Forum for Information and Documentation* 17(2), 1992.

A. Knott. "New strategies and constraints in RST-based text planning." M.Sc. thesis, Dept. of Artificial Intelligence, University of Edinburgh, 1991.

L. Kosseim, G. Lapalme. "Choosing rhetorical relations in instructional text: the case of effects and guidances." In: Proceedings of the 5th European Workshop on Natural Language Generation, Leiden, 1995.

W. Mann, S. Thompson. "Rhetorical structure theory: Towards a functional theory of text organization." In: *TEXT*, 8:243-281, 1988

D. Marcu. "The rhetorical parsing, summarization, and generation of natural language texts." Ph.D. thesis, Dept. of Computer Science, University of Toronto; Technical report CSRG-371, 1997.

C. Matthiessen, J. Bateman. *Text generation and systemic-functional linguistics*. London: Pinter, 1991.

K. McKeown. *Text Generation.* Cambridge University Press, 1985.

J. Moore, C. Paris. "Planning text for advisory dialogues: Capturing intentional and rhetorical information." In: *Computational Linguistics* 19(4), 1993.

C. Paris, K. Vander Linden, M. Fischer, A. Hartley, L. Pemberton, R. Power, D. Scott. "A support tool for writing multilingual instructions." In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95), Montreal, 1995.

W. Ramm, C. Villiger. "Global text organization and sentence-grammatical realization." CLAUS report 61, Universität des Saarlandes, 1995.

D. Rösner, M. Stede. "Customizing RST for the Automatic Production of Technical Manuals." In R. Dale et al. (eds.) *Aspects of Automated Natural Language Generation.* Berlin: Springer, 199-214, 1992.

D. Scott, C. de Souza. "Getting the Message across in RST-based Text Generation." In: R. Dale et al. (eds.) *Current Research in Natural Language Generation.* Academic Press, 31-56, 1990.

C. Umbach, M. Stede. "Kohärenzrelationen: Ein Vergleich von Kontrast und Konzession." Submitted.

K. Vander Linden, J. Martin. "Expressing rhetorical relations in instructional text." In: *Computational Linguistics* 21(1):29–58, 1995.

J. Williams. *Style. Towards clarity and grace.* University of Chicago Press, 1990.

L. Wanner. "On lexically biased discourse organization in text generation." In: Proceedings of COLING-94, Kyoto, 1994.