Res Lang Comput (2008) 6:311–332
DOI 10.1007/s11168-008-9053-7

Disambiguating Rhetorical Structure

Manfred Stede

Published online: 30 December 2008

Abstract Empirical studies of text coherence often use tree-like structures in the spirit of Rhetorical Structure Theory (RST) as representational device. This paper identifies several sources of ambiguity in RST-inspired trees and argues that such structures are therefore not as explanatory as a text representation should be. As an alternative, an approach toward multi-level annotation (MLA) of texts is proposed, which separates the information into distinct levels of representation, in particular: referential structure, thematic structure, conjunctive relations, and intentional structure. Levels are conceptually built upon each other, and human annotators can produce them using a dedicated software environment. We argue that the resulting multi-level corpora are descriptively more adequate, and as a resource are more useful than RST-style treebanks.

Keywords Discourse · Coherence · Rhetorical structure · Multi-level analysis · Corpus annotation

1 Introduction

Discourse researchers largely agree that *coherence relations* are an important instrument for describing the "information surplus" that distinguishes a text from a mere sequence of sentences: The relations characterize the information communicated by the writer through juxtaposing two text segments, in addition to the segments considered in isolation. These phenomena have been studied from the viewpoints of formal semantics (e.g., Asher and Lascarides 2003; Kehler 2002), cognitive modelling (e.g., Sanders et al. 1992), extending syntactic description to the discourse level (e.g., Polanyi

M. Stede (✉)
Institute of Linguistics, University of Potsdam, Potsdam, Germany
e-mail: stede@ling.uni-potsdam.de

1988; Webber et al. 2003), or empirical observation (e.g., Mann and Thompson 1988; Carlson et al. 2003; Wolf and Gibson 2005).

The viewpoint of the present paper is characterized by the goal to explain coherence by studying authentic texts, using annotated corpora as a supportive tool. An influential account along these lines was the proposal of Rhetorical Structure Theory (RST; Mann and Thompson 1988; Matthiessen and Thompson 1988; Taboada and Mann 2006), which later lead to an annotated corpus (Carlson et al. 2003). We take this work as a starting point, but in contrast to these authors, we question the descriptive adequacy of a single tree structure for capturing discourse coherence and, consequently, we question the utility of corpora annotated with such trees. In particular, we argue that rhetorical trees encapsulate diverse kinds of information, which ought to be kept separate in an explanatory text representation. When producing an analysis along the lines of RST, one often has to make decisions on the grounds of judgements that (i) pertain to different levels of description, and (ii) are not explicitly recorded in the resulting overall tree structure, which leads to considerable ambiguity in the analysis process and the resulting trees. For illustration, consider this example:

(1) (a) We saw a two-months-old polar bear cub at the zoo yesterday. (b) It kept climbing on top of the keeper's shoulders all the time. (c) With its teddy-like head it was the cutest thing I've seen in years!

One analysis along the lines of RST, warranted both by the relation definitions from Mann and Thompson (1988) and the annotation guidelines of Carlson and Marcu (2001), treats (b) and (c) as individual *Elaborations* of (a), as they provide additional information on the entity introduced by (a). By definition, (a) becomes the only nucleus of the resulting structure.[1]

However, (a) can also be analyzed as satellite of a *Background* or *Circumstance* relation to (b) and (c), which in a way are the "more important" pieces of information and which can internally be linked by *List* or by *Conjunction*. As for (a), both Mann/Thompson and Carlson/Marcu would probably give a slight preference to *Circumstance*, but—interestingly—for decidedly different reasons: Mann/Thompson see *Background* as a "presentational" relation, thus emphasizing the intentional role of the segment (satellite increases reader's ability to comprehend an element in the nucleus), and *Circumstance* as a "subject-matter" one, according to which the satellite "sets a framework in the subject matter within which the reader is intended to interpret the nucleus". Arguably, in our example the subject matter plays a somewhat greater role than the intentions. Carlson and Marcu, on the other hand, suggest that the analyst prefer *Circumstance* when the events described in nucleus and satellite are "somewhat co-temporal", which is certainly the case in our example. Note that these criteria employed by Mann/Thompson and Carlson/Marcu operate on different levels and can therefore easily be in conflict with one another in a given context.

Another candidate relation for the text is *Topic-Comment*, defined by Carlson/Marcu as multinuclear. The first element [here: (a)] introduces a topic of discussion, and the

[1] Carlson and Marcu (p. 36) propose a deletion test for deciding on the structure of elaborations: Since the text would be acceptable when segment (b) were missing, both (b) and (c) are independent satellites, rather than (c) elaborating (b) and the two collectively elaborating (a).

second [here: (b) and (c)] makes a statement about this topic; this is also a valid description of our example. And, to give one more possible reading, the analyst might be prompted by "the cutest thing" in (c) to record that the author expresses his or her personal opinion; the right relation in this case is *Evaluation*. Following Mann/Thompson, (c) is by definition the satellite and (b) the nucleus, as the general description in (a) is maybe not being found cute—only the animal is. Carlson/Marcu, on the other hand, offer three variants of *Evaluation*: Either the evaluating segment or the evaluated one or both can be nuclear. We leave it to the reader to decide which one might be most appropriate in the example.

The supposition that rhetorical structure analysis be a task involving ambiguity might not come as a surprise to the RST analyst, since Mann and Thompson (1988) themselves had already pointed out that, quite naturally, different analysists will at times make different decisions. After all, constructing an RST tree for a text involves *interpreting* the text, i.e., reconstructing the various intentions of the writer, which is bound to be subjective: My interpretation is different from your interpretation, we take different messages away from the text, and by inspecting the RST trees, we can to some extent track the reasons for our arriving at different conclusions down to different decisions on relations, their nuclei, and the spans of text they relate.

The situation is different, however, when ambiguity becomes a problem also within the work of a single analyst. When he or she repeatedly cannot see how to make a decision between alternatives (of relations, spans, or nuclearity assignment) on principled grounds, then at the end of the overall process the analyst has produced an RST tree but finds that the result might just as well be one or several different ones. This will leave a good analyst frustrated, because what is the point of drawing "the" tree when it involves a good deal of arbitrariness? One way of responding to this situation is to devise more explicit, and more complex, annotation guidelines. This path was pursued by Carlson et al. (2003) with their annotation manual of the RST Treebank corpus. They use 85 relations and hence make more distinctions than those in the original RST proposal (24 relations). But the overall annotation task does not become any easier: Analysts now need to have the portfolio of 85 relations at least roughly in mind (for they need to know when and where to consider and look up a relation definition), and with that many relations, the number of alternative analyses of even a short text is bound to increase (cf. our discussion of example 1 above). Furthermore, the *structural* problems associated with RST analysis (see Sects. 2.1–2.6 below) stay the same, irrespective of the number of coherence relations.

After considering the sources of ambiguity in RST-style trees in more detail (Sect. 2), this paper proposes to decompose discourse structure into four separate levels, which belong to different conceptual realms, and which *collectively* render a text coherent. Thus Sect. 3 introduces the framework of multi-level text analysis (MLA), which aims at "deconstructing" discourse structure into a number of less complex levels of description. MLA is being implemented with the *Potsdam Commentary Corpus*, henceforth *PCC* (Stede 2004), and Sect. 4 briefly describes the software environment we developed for annotating and maintaining MLA corpora, and the annotation practice we use with the PCC. Finally, Sect. 5 compares MLA to related approaches, and Sect. 6 summarizes the central points of the paper.

2 Ambiguity in Rhetorical Trees

RST, as it was pointed out by Knott et al. (2001), is a theory both of text organization (hierarchy; a general nuclearity principle; tree structure rather than general graph) and of specific coherence relations and their definitions. Analysts producing an RST tree thus have to make a wide range of different decisions: What is a minimal unit of analysis? Which adjacent units are to be grouped together by a relation? Which relation? Which unit is more important (nuclear) than its neighbour? Where are the boundaries of larger text units and how do they relate to one another? What is the difference in importance between those larger units? What units contribute most centrally to the overall purpose of the text? How does this propagate down to the nuclearity assignment at the level of minimal units [cf. the "strong nuclearity principle" of Marcu (2000)]? The problem is not so much that analysts be unable to cope with all these decisions, but that the answers they gave to the manifold individual questions are in the end only partially visible in the tree. Many will have indirect effects that contribute to the overall structure but are not explicitly represented, thus leaving the structure ambiguous.

2.1 Intention or Information

The one ambiguity that has been widely discussed in the RST community is that between an analysis in terms of *presentational* (or *intentional*) and *subject-matter* (or *informational*) relations. Moore and Pollack (1992) showed examples that invite analyses on either level, and that could even lead to opposite assignments of nuclearity:

(2) (a) Come home by 5 o'clock. (b) Then we can go to the hardware store before it closes.

This sentence pair can be analyzed as subject-matter *Condition* ("if you come home early, then we can go to the store") and as presentational *Enablement* ("your coming home early is a prerequisite for visiting the store"), and in both cases (b) is the nucleus of the relation. At the same time, an equally plausible reading is a *Motivation* relation ("we want to go to the store, so make sure you come home early"), in which case (a) is the nucleus.

The consequences of observing such examples are much less clear. Should every text in principle be analyzed on both the presentational and the informational level? In our experience with the PCC (which originally we had annotated according to RST), relation ambiguity is a problem, but the Moore/Pollack type of *systematic* ambiguity between informational and presentational analyses does not occur often enough as to generally warrant a complete analysis on both levels; most of the time, annotators are not torn between assigning an intentional and an informational relation.[2] The

[2] Example 2 also points to the fact that the division between subject-matter and presentational relations as proposed by Mann and Thompson is a bit problematic; both paraphrases refer to the prospective action of the hearer, so why should the *Condition* reading be on the 'subject-matter' level? Also, for example 1 we mentioned *Evaluation* as a possible relation; Mann and Thompson regard it as subject-matter, but it clearly involves a subjective attitude of the author, which is different from merely describing (complex) states of affairs in the world.

same point was made by Sanders and Spooren (1999), who argued against assuming a complete dual information/intention analysis. Therefore, if we record both types of structure, they will often be sparse: In narrative or descriptive text, not much might happen on the intentional side, whereas in argumentative text, the informational structure is more likely to have gaps.

2.2 Speech Act and Epistemic Status

So far, theories of discourse pragmatics have centered their attention either on the role of speech acts (in the Austin/Searle tradition) as one-place functions, or on two-place coherence relations—but to our knowledge, no theory sees the need for accommodating both. Some relation definitions by Mann and Thompson, however, point to differences in "status" of the related units by constraining the types of nucleus and satellite. When inspecting these constraints from all their relation definitions, it turns out that they can be roughly divided into more semantic and more pragmatic features:

- Semantic: not unrealized; hypothetical, future or otherwise unrealized situation; activity; volitional action; unrealized action in which reader is actor.
- Pragmatic: writer has positive regard for X; X presents a problem; reader won't comprehend X before reading Y; reader might not believe X; reader believes X or will find it credible.

These features, however, are not defined more thoroughly in RST, let alone organized in any system. Nor do *all* relation definitions make use of such constraints on the related units. As a result, when an analyst assigns a relation, this sometimes indirectly implies a statement on the underlying speech act and/or the epistemic status of the units, and sometimes it does not. Consider the following example:

(3) (a) The museum has closed already. (b) Most lights are off.

If (a) is read as a factual statement made by the writer, then (b) is in an *Elaboration* relationship to it: The topic 'museum' is maintained, attention is focused on one particular aspect of it. However, in a context where the conversants are debating whether they should visit the museum, (a) can be read as an estimate or claim, for which (b) provides substantiation; this would be a clear case of an RST *Evidence* relation. From a semantic viewpoint, the difference corresponds to that between presupposed and asserted information; from a pragmatic viewpoint, this has been encoded in inventories of *illocutions* as more fine-grained versions of Searle's (1976) taxonomy of speech act types. We will in Sect. 3.5 use the inventory suggested by Schmitt (2000) as labels for minimal discourse units that *in conjunction with* a coherence relation provide a more explicit and more complete picture of the analyst's interpretation than an RST tree, whose relations only sometimes indirectly refer to illocution types.

2.3 Units of Analysis: What Exactly is being Related?

A quite critical issue for assigning coherence relations in authentic text is that of defining the minimal units of analysis. Mann and Thompson state that these be "typically clauses" and do not elaborate this in much detail. Carlson and Marcu (2001) are

more explicit here, but still leave a number of questions open. Consider the following example:

(4) (a) Yesterday at Cecilienhof Palace, the local business people lamented that they will have to lay off quite a few employees.
(b1) Thus unemployment in the area will continue to rise.
(b2) Thus in our town litany follows upon litany.
(b3) As usual, Cecilienhof was beautifully decorated for the meeting.

The sequence (a–b1) displays a type of ambiguity that our annotators of the PCC encountered quite often. It can be analyzed as (a) being either the *Nonvolitional-Cause* or *Evidence* for (b1). Notice that here the analysis depends on the choice of unit: The *Evidence* reading relates (b1) to the complete lamenting event in (a), while the *Cause* reading relates it only to the embedded clause: Rising unemployment is not caused by the business people pointing something out but by their laying off employees. Sentence (b2), using the same connective, can only link to the complete (a), i.e., to the lamenting, with a suitable relation being *Evaluation*. (b3), finally, can be analyzed as *Elaboration*. This, however, is a link from (b3) neither to the embedded laying-off event nor to the lamenting event. Instead, it elaborates on the location expressed by an adjunct PP in (a).

Deciding on which RST relation to assign thus depends on which portion of (a) is selected as the relation's argument, and this sub-decision will not be transparent in the resulting tree. A part of this problem had been addressed by Carlson and Marcu (2001), who treat the embedded *that*-clause in (a) as a separate minimal unit, linked to the matrix clause by a relation called *Attribution*. This indeed handles many cases in the genre they analyzed (business news). Still, it is not clear how to connect the (b) segment to just one of the units within (a), unless one subscribes to the "strong nuclearity principle" and also allows *Attribution* to come in two nucleus-satellite assignment variants [which Carlson/Marcu did not choose to do; for a critical discussion, see Redeker and Egg (2006)]. But even then, important consequences remain the same: In contrast to "standard" relations holding between complete eventualities, splitting up a sentence like *The guard said that the museum was closed* leads to fairly artificial semantic objects corresponding to the units "The guard said" or "The guard said that". *Attribution* thus does not have the same status as, say, relations of causality or contrast: The relationship between an event of saying and the specific contents of that saying is different from a coherence relation linking two complete propositions. Furthermore, the problem of embedded clauses extends beyond verbs of communication and hence beyond the coverage of *Attribution*. Consider "The business people finally decided to lay off quite a few employees", which also can be followed by a segment that links either to the laying-off event or to the deciding event.

2.4 Enforced Nuclearity

With the exception of the three multinuclear relations and the Cause/Result relation pair, which offers opposite assignments of nucleus and satellite to the related text spans, choosing an RST relation in the (Mann and Thompson 1988) framework automatically involves assigning nuclearity status to one of the spans. Conversely, judging

one segment as nuclear reduces the range of admissible relations (to those that render precisely that segment nuclear). Carlson and Marcu (2001) saw this as too strict a limitation and responded by introducing a variant with opposite nucleus/satellite assignment for many relations. In our experience, however, there are many contexts where annotators feel uncomfortable having to select one nucleus at all. Consider example 4(a–b1) above, which might occur in a news message where the context does not particularly emphasize either of the two segments. Then, the nucleus decision is a rather arbitrary one, and thus a discourse representation should offer the chance to not mark any nucleus, rather than enforcing a choice.

Furthermore, even in situations where different annotators easily agree on which unit should be labelled nuclear, the *reasons* for identifying a nucleus can vary considerably, and the decisive reason is not recorded in the representation. This is discussed in detail in Stede (2008). For our purposes here, we focus on the ambiguity that can arise from the fixed association between relations and participating nucleus and satellite. The RST-analyst quite often is given a choice between two relations whose definitions both apply to the segments in question, but which assign nuclearity in opposite ways. Consider example 5, which offers some of the ambiguities we already encountered in our introductory example 1. Segment (a) can be interpreted as providing *Background* for (b) (the nucleus), or (b) can be seen as an *Elaboration* of (a) (the nucleus).

(5) (a) Yesterday I bought a new Buick. (b) For a test, I drove it all the way to Salt Lake City.

The analyst can make a decision either by carefully weighing the subtleties of the relation definitions, or on the grounds that she wishes to render either A or B nuclear, for reasons imposed by the context. This has been observed already by Bateman and Rondhuis (1997) who regard it as an advantage that global nuclearity considerations can assist in the local decisions with assigning a relation. From the perspective of ambiguity, however, the problem (indeterminacy as to the relation) is not really resolved but hidden: The final RST tree does not indicate whether some relation at the level of minimal units is there because its definition is optimally fulfilled or because text-global factors make it seem advantageous to select one particular nucleus, which is incidentally performed by that particular relation.

2.5 Scope of Relations

When a subordinating or coordinating conjunction is present in the text, the boundaries of the spans of the coherence relations are usually clear. But with adverbial connectives or no lexical signal at all, the scope is sometimes quite difficult to identify. PCC annotators consistently report difficulties with deciding on spans higher up in the tree. Consider the following excerpt (translated into English) from one of the PCC texts:

(6) In our city there are very few apartments that are at the same time located in nice areas, affordable, and available. Especially for people with low income, the market is extemely tight. (…) This leads to the trend that many citizens are being forced to move to less attractive areas, because only there they can find small and

inexpensive apartments. Thus in the long term, the city is being divided even more dramatically into "rich" and "poor" areas. Not a nice development.

The last sentence is quite clearly the satellite of an Evaluation relation, but what exactly is the nucleus? Is it only the penultimate sentence or also one or two of its predecessors, or even the entire text? This is the type of decision that cannot confidently be made by considering just the relation definition; instead, it is governed by the need to produce an overall spanning discourse tree. Making these decisions often leads to disagreement among annotators: The longer the text, the more difficult it becomes to keep track of the different portions and to join them in the "right" hierarchy on principled grounds.

2.6 Vague Relation Definitions

To conclude this overview of sources of ambiguity in rhetorical trees, we have to mention the problem of vagueness in relation definitions as they were provided by Mann and Thompson (1988). This has been criticized in the literature before, so we can be fairly brief. Consider as one example the definitions of *Antithesis* and *Concession*. The constraints on the nucleus and the intentions of the writer (i.e., the "effect") are identical. *Antithesis* has no constraint on the satellite, whereas *Concession* offers the constraint that "writer is not claiming that satellite does not hold". (Since *Antithesis* has no constraint here, does it properly subsume *Concession*?) Finally, the constraints on the nucleus/satellite combinations are largely paraphrastic with the one exception that Antithesis adds that "one cannot have positive regard for both situations" (in nucleus and satellite). In total, the differences are not very restrictive, so that in many contexts both definitions are equally applicable. But, in the presentational/subject-division of the relations suggested by Mann and Thompson, *Antithesis* appears in the former, and *Concession* in the latter, despite their effects being identical. So it is not clear on what grounds the grouping is made in this case.

Returning again to example 1, we note that the definitions of *Elaboration, Background* and *Circumstance* are so imprecise that they are simultaneously applicable in a great many cases. As a response, many annotators seem to resort to *Elaboration* as a "default", so that the presence of this relation in some RST tree can result either from the spans being in a genuine, "good" Elaboration relation (and the annotator confidently assigned it) or from a perceived unclear relationship between the spans, which is somehow also covered by *Elaboration*.

To some extent, it is probably inevitable to pay the price of relation vagueness when proposing a set of coherence relations designed to cover a great variety of authentic texts (cf. the remarks at the beginning of Sect. 1). But the problem is aggravated by RST's taking the step to join both 'presentational' and 'subject-matter' relations in the same framework, which on the one hand involves definitions that appeal to author's intentions and presumptions on reader's mental states, and on the other hand leads to the problem of overlap mentioned in Sect. 2.1. In our proposal described in the next section, we therefore suggest to move the intentional relationships to a separate level of description, and to reduce the role of other coherence relations to a rather "flat", surface-oriented representation.

2.7 Interim Summary

In this section, we have identified a range of problems with RST-style analysis of texts, which can be summarized as follows.

- P1 (Sect. 2.1): Discourse segments cannot be simultaneously related on the informational and the intentional level of description.
- P2 (Sect. 2.2): Assumptions of the analyst on the speech act and epistemic status of a segment are not made explicit (but they can have consequences for the rhetorical tree).
- P3 (Sect. 2.3): Assigning a relation can depend on the portion of a segment that the analyst focuses on—which is not made explicit.
- P4 (Sect. 2.4): Sometimes, there are no good reasons for choosing one of two related segments as nucleus—but the theory enforces a choice.
- P5 (Sect. 2.4): Sometimes, there *are* good reasons for choosing a nucleus—but the definition of the appropriate relation contradicts them.
- P6 (Sect. 2.4): There may be quite different reasons for assigning a nucleus, but these are not made explicit in the tree.
- P7 (Sect. 2.5): Sometimes, it is so difficult to decide on the scope of a relation that the choice is basically arbitrary.
- P8 (Sect. 2.6): Due to vague relation definitions that often operate on different levels of description, quite often more than one relation is applicable to a pair of segments.

3 Discourse Structure Revisited: Multi-Level Analysis

3.1 Overview

We have identified a number of ambiguities and annotation problems with Rhetorical Structure Theory, and our proposal in this section is that in essence, they stem from the desire to capture the coherence of a text within a single tree structure. As an alternative, we view coherence as resulting from the interplay of different levels of description, to be captured in a framework of *multi-level analysis*, henceforth *MLA*. The four levels proposed here are:

- Referential structure: Co-reference relations
- Thematic structure: Hierarchical structure showing (sub-)topic shifts
- Conjunctive relations: Surface-oriented coherence relations (Martin 1992)
- Intentional structure: "Deep" support relations between text segments and their illocutions

Coreference is traditionally seen as a central aspect of text coherence, and we wish to reinforce the view that it needs to be accounted for in addition to relation-induced coherence. *Thematic structure* records what the text is about and what topics are subsumed by others. And for a coherent text, it should be possible to name the topic of the entire text, with no segment being "lost" (cf. van Dijk 1977). *Conjunctive relations* are links that can be read off the text surface without performing "deep" inferences; these relations can be directed but they do not assign different degrees of prominence

to the relata. Crucially, in MLA it is also possible that adjacent text segments are not linked by any such relation. *Intentional structure* consists on the one hand of illocution labels for the minimal units, and on the other hand of a tree structure whose precise shape depends on the text type. We consider here only argumentative text, where links can either indicate simple "support" relations, or build potential counter-arguments or their dismissal. This structure is also incomplete: Not every segment needs to contribute directly to the unfolding argument—it might merely serve to mark a topic shift or constitute some less relevant rhetorical move.

Before explaining the levels in more detail, we briefly compare this approach to RST-like representations. In contrast to these, the only level of analysis that MLA requires to span the entire text is thematic structure. Both conjunctive relations and intentional structure will usually each cover only parts of the text. Another important deviation from an RST principle is the possibility to link non-adjacent segments in argument structure. This in in accord with the proposal of an 'illocution structure' by Brandt and Rosengren (1992): Support relationships between illocutions are not limited to adjacent units; the structure of an argument can be more complex than a tree structure suggests.

By distributing relational information to the two levels of intentional and conjunctive relations, MLA responds to problem P1 identified in Sect. 2.7: Two segments can simultaneously be in an intentional support relation and in a "surfacy" conjunctive relation. P2 is addressed on the level of intentional structure, which records the illocutions of segments so that this information transparently complements the relation assignments.

The problems with nuclearity (P4–P6) are handled by abolishing the nucleus/satellit distinction as a general principle underlying a single rhetorical tree. Instead, different degrees of prominence on the intentional level are encoded by the support relations, while other differences in prominence can be captured by thematic structure, which records beginnings and endings of topic zones. Also, notice that MLA does not recognize a relation corresponding to RST's *Elaboration*. (This argument has in similar ways been made by Knott et al. (2001) and Poesio et al. (2004)). Instead, in MLA such a relationship will typically be a combination of a coreference link and a continuing topic. This might also co-occur with a conjunctive relation; for example, it is not difficult to state a *Contrast* that also qualifies in RST terms as *Elaboration*: "The student's performance generally received good marks. Only his oral presentations were frequently criticized." In RST (or any other single-level account), only one of the two relations can be accounted for.

Future work will add a level of sentential information structure as a mediator between syntax and discourse level (in particular thematic structure). By introducing topic and focus markings, another underlying source of RST-type nuclearity will be made transparent. And this will also contribute to an answer to P3: If the focus of a segment is marked, it can be set into correspondence with the intentional or conjunctive relation the segment participates in. Note, however, that the possibility to assign both kinds of relations already partially resolves P3, as a segment can be in two relations to its environment, on different levels of description.

P7 is addressed by enforcing precise segment delimitation (beyond minimal units) only on the level of intentional structure. Here, for the text type we consider, tracking

Mandatory vaccination against children's diseases?
[1] Today, children don't know anymore what pox are. [2] What a joy. [3] When pox vaccination was introduced in 1854, [4] quite a few people believed [5] that their head would turn into a cow's head [6] if they got themselves vaccinated. [7] For the vaccine was made from cattle's skin at the times. [8] Nowadays this dreadful disease is exterminated. [9] Thanks to a determined, world-wide vaccination campaign. [10] But there still are other diseases: measles, polio, diphtheria, mumps, rubella, hepatitis B, tuberculosis, pertussis. [11] Millions of children die of these, especially in less developed countries. [12] In Germany, many parents apparently don't take these diseases seriously. [13] Because they don't know them anymore! [14] For it has been achieved with vaccines [15] that these infections hit only rarely today. [16] But those who have experienced [17] how terribly children suffer [18] when they come down with 'just' measles or pertussis, [19] should spare them the agony. [20] As well as the long-term consequences. [21] Only those who have their children vaccinated will contribute to vaccines' becoming superflous some day. [22] Instead, people rant about side effects [23] that occur very rarely and are known merely from books. [24] Then there is the great argument: This is my child, the government must not prick her. [25] No vaccine can help against such parents.

Fig. 1 Sample text "vaccination"

down the argument relies on identifying quite precisely how statements are intended by the author to support one another. In addition, larger segments arise on the level of thematic structure, where they are identified by topic switches; these do not need to correspond to relation-induced segments, which we have identified as a central difficulty in RST. P9, finally, is approached by separating the set of relations into the more surface-oriented conjunctive ones and the decidedly "deep" intentional ones. This allows for defining relations within their respective realm, so that they can be set apart from each other *within* that realm more easily, and furthermore, *between* the realms the definitions need not exclude each other anymore (which RST aspired to in order to produce a single tree).

The following subsections introduce the levels of MLA in more detail. We illustrate the approach with a text from the PCC, shown in Fig. 1 in an English translation. It is designed to preserve the German clause structure and connective usage, and thus it might not sound perfectly "natural" in a few places. Segment numbers have been inserted by the author.

3.2 Coreference

An explanation of a text's coherence cannot be complete without making the coreference relations explicit. Referential continuity is not only a defining feature of "textuality" [cf. the notion of 'entity coherence' put forward by Knott et al. (2001) and Poesio et al. (2004)]; it also contributes to other levels of analysis. The thematic structure (see below) is informed by the development of referential (dis-)continuity, and in some cases the assignment of coherence relations can depend on what the analyst regards as coreferential or not, e.g., in deciding whether there is some implicit causal relation between reported events.

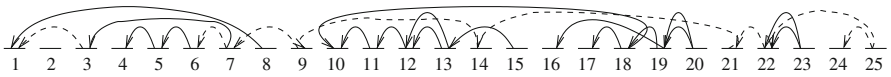

Fig. 2 Referential structure of sample text "vaccination"

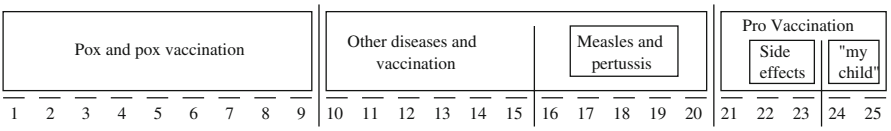

Fig. 3 Thematic structure of sample text "vaccination"

Figure 2 shows an analysis of the sample text from Fig. 1, giving only a coarse-grained segment-based view, which is sufficient for our purposes here. Solid arrows represent an anaphoric link between two segments; dashed arrows indicate a bridging anaphor, as for example that between *pox vaccination* (3) and *pox* (1). Besides showing only the segments rather than the referring phrases, the figure also omits the information on identity of referents, i.e., the distinction between separate chains.

3.3 Thematic Structure

This level of description tracks the thematic development of the text and bears some similarity to the "attentional structure" of Grosz and Sidner (1986). One purpose is to identify text segments that contain formulaic text portions (opening, closing, greeting, etc.) or meta-discursive elements like "In this section, we will show that…". These do not drive the content of the text forward but provide text-internal information, either schematic or provided by the author to orient the reader. Units of this kind receive the label *conventional*; our sample text does not contain such a unit.

The more important purpose is to break the text into segments that address the same topic, and to indicate sub-topic relationships where appropriate. Figure 3 gives an analysis along these lines for the sample text. The notation uses vertical lines to indicate topic changes (the first step of the analysis) and boxes with short descriptions of the topic for segments.[3] Boxes can be embedded in one another to indicate sub-topic relationships.

A closer look at Figs. 2 and 3 reveals certain similarities. In fact it had already been suggested by Figge (1971) that the "referential chains" of a text quite often provide useful hints as to its thematic structure. Intuitively, this is not suprising: If the topic changes, some or most or all of the entities being referred to should also change. For our sample text, indeed the vertical lines in the thematic structure occur at weakly-connected portions of the referential structure. The point here is that analyzing and representing these levels of information separately allows for studying the relationships between these linguistic domains more thoroughly, as well as interactions with

[3] The form of a felicitous topic description (word, phrase, sentence, question) is to a good extent dependent on the text type; see the extensive discussion of Lötscher (1987).

other levels (such as the chosen types of referring expressions, or linguistic devices for signalling topic change, for instance).

3.4 Conjunctive Relations

For our level of coherence relations, we use the work of Martin (1992), who provided a thorough analysis of (English) connectives and the relations they can express. His *conjunctive relations* are decidedly surface-oriented and reflect only features of linguistic realization plus minimal semantic interpretation, but no judgements on underlying speaker intentions. The relations can be *directed* to show a semantic distinction among the elements (as is obvious for causal relations or for *Condition*), but the elements are of equal pragmatic "importance", i.e., there is no notion corresponding to nuclearity. Martin offers a very fine-grained classification of relations and ways of their linguistic realization; we are using here only the 12 "basic" relations (Martin 1992, p. 179), which are categorized in four groups:

- Temporal: *Simultaneous, Successive*
- Consequential: *Manner, Consequence, Condition, Purpose, Concession*
- Comparative: *Similarity, Contrast, Reformulation*
- Additive: *Addition, Alternation*

Martin emphasizes that conjunctive relations have both an "internal" and an "external" reading. The former operate within the text and reflect its organization, while the latter mirror relationships between eventualities in the world. Accordingly, Martin divides the set of linguistic connectives into those that are distinctly internal and those that can be both internal and external. The latter in turn are grouped into "cohesive" connectives (essentially adverbials), coordinating conjunctions and subordinating conjunctions. For example, a distinctively internal signal of the *Contrast* relation is "on the other hand"; its internal/external signals include "in contrast" and "instead" (cohesive), "but" (coordinating), "whereas" and "except that" (subordinating). In the sample text, a clear example of an internally-used connective is *then* in (24), which expresses not a temporal relation between eventualities in the world, but the order of presentation of the author's arguments.

In a step that departs from the close association with the linguistic surface, Martin acknowledges the existence of "implicit connectives" in text, that is, relations that are not signalled by a connective. The analyst is encouraged to assume a relation when some connective signalling that relation could be present in the text without altering the perceived meaning. Frequent cases are temporal *Succession*, which is a default relationship when eventualities are narrated in a text and thus no signal need be present; *Addition* for listings of related observations; *Contrast* when it is signalled by lexical means other than a connective. In the sample text, this is the case between segments (11) and (12): "in less developed countries/in Germany". In order to avoid ambiguity in assigning hierarchical structure, we restrict the application of implicit relations to adjacent *minimal* units only. Quite often, one of the aforementioned relations is applicable when no connective is present in the text; but this is not always the case. Between units (1) and (2), no connective can be construed, not even the quite general *Addition* is applicable here. Figure 4 shows an analysis of the text in terms of

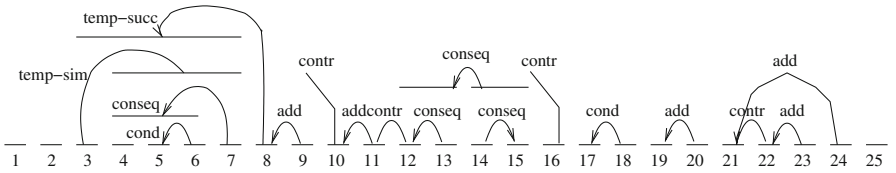

Fig. 4 Conjunctive relation structure of sample text "vaccination"

Martin's conjunctive relations. Directed relations are marked with an arrow. Notice that larger segments are formed only when the scope of a connective is very clear, as with the subordinating *when* in (3–7). The analysis also makes use of a final feature of Martin's scheme, which is very different from RST: In case the scope of a connective is not clearly recognizable, the analysis can leave it unspecified. This holds for the contrastive connectives in (10) and (16). In the diagram, the direction of the line indicates that the left boundary of the segment is left open.

3.5 Intentional Structure

While conjunctive relations reflect judgements that apply on a local level and are often prompted by surface-linguistic signals, the "deep" level of analysis involves the reconstruction of speaker intentions. By definition, judgements on this level are more subjective, because the analyst has to interpret the text and state *why* (in his or her view) the author wrote a particular portion of the text and how it relates to other portions.

To this end, we begin the intentional analysis with a level of *illocution* or *speech-act* assignments. That is, the text is segmented into units the analyst considers as performing an individual speech act, and each such unit receives a label for its underlying illocutionary force. An inventory of such labels has been suggested by Searle (1976), and an extension was proposed by Schmitt (2000). As Schmitt's inventory deals with both the more "traditional" intention-oriented illocutions and with different epistemic stances the speaker can take towards the proposition, it is particularly well-suited to reflect distinctions in argumentative text, where it is central to make claims and substantiate them with statements that are (supposedly) beyond dispute. A crucial aspect of understanding such a text thus consists in identifying the epistemic and, more generally, pragmatic status of the units. The relevant subset of Schmitt's categories of illocutions are:

– *Reportivum*: Writer describes a state of affairs.
– *Identifikativum*: Writer characterizes his or her own state of mind, health, etc.
– *Estimativum*: Writer asserts proposition as *probably* true.
– *Evaluativum*: Writer presents a personal opinion.
– *Appellativum*: Writer orders or suggests an action.

The units of analysis can be the minimal units defined in the earlier phase of analysis, or combinations of adjacent units. In our sample text, (1) and (2) are a complex *evaluativum* (as the elliptical (2) cannot in itself be assigned an illocution). The same holds for (8–9). Usually, subordinate clauses do not constitute separate illocutions;

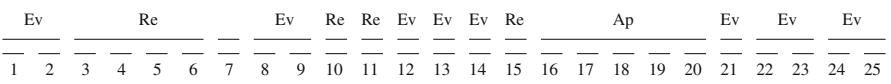

Fig. 5 Illocution analysis of sample text "vaccination"

thus (3–6) form a single illocutionary unit. But since our labels also pertain to epistemic status, there are exceptions. Consider (14–15): A plausible analysis sees (15) as a *reportivum*, since the low figures of infection cases nowadays are common wisdom. Whether this development is due to vaccination, though, could very well be disputed; this is merely a hypothesis. In this light, (14), which proposes a reason for (15), should be analyzed as *evaluativum*. A complete assignment of illocutions to units of the text is shown in Fig. 5.

The individual illocutions now serve as building blocks for the intentional structure of the text. Thus we subscribe to the view of Sanders et al. (1992) (and others), who distinguish between *semantic* coherence relations, which hold between propositions, and *pragmatic* relations, which hold between speech acts. In line with Grosz and Sidner (1986), Brandt and Rosengren (1992), Moser and Moore (1995) and others, we take the intentional structure to be a graph in which a directed *dominance* (or, from the opposite perspective, *support*) relation can hold between (the illocutions of) two elementary discourse units or recursively between larger units. But we do not assume that such a graph has to span the entire text—there need not always be an intention-based connection between adjacent units of text. Further, while Grosz and Sidner deny the possibility of systematically distinguishing subtypes of this relation, we follow Brandt and Rosengren and also Mann and Thompson (1988) in their attempts to devise such a classification and use the following types:[4]

1. *Ease-understanding*: Supporting speech act enables reader to understand the propositional content of the supported speech act (in RST: Background)
2. *Encourage-acting*: Supporting speech act encourages reader to perform the action stated in the supported speech act (in RST: Motivation)
3. *Ease-acting*: Supporting speech act enables reader to perform the action stated in the supported speech act (in RST: Enablement)
4. *Encourage-believing*: Supporting speech act encourages reader to believe the content of the supported speech act (in RST: Evidence)
5. *Encourage-appreciating*: Supporting speech act encourages reader to develop a more positive regard towards the content of the supported speech act (in RST: Antithesis, Concession)

For the purposes of this paper, however, we focus specifically on the type of text that a part of the PCC deals with: short argumentative pieces that typically respond to a "Should we do X" question. For these, the "endpoint" of understanding is a reconstruction of the argument the author is presenting, which can be seen as a specialization of the

[4] This list generalizes over the 'presentational' relations given in Mann and Thompson (1988). On the RST website (www.sfu.ca/rst), several other relations are classified as 'presentational' (e.g., Summary, Restatement), but we do not agree that they should be treated *on a par*, as there is no "supporting" to be discerned.

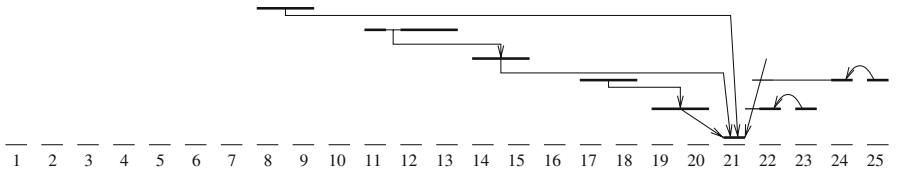

Fig. 6 Argument structure of sample text "vaccination"

general intentional structure. It includes primarily the relations *encourage-believing* and *encourage-appreciating* from the above list, but we also have to provide for the refutation of possible counter-arguments. In the sample text, two counter-arguments are mentioned (briefly), in segments (22) and (24), and they are immediately dismissed by the author in (23) and (25). A useful notation for capturing the structure of argumentative text has been proposed by Freeman (1991), who built upon the well-known scheme by Toulmin (1958). While the latter is a static instrument to represent arguments-as-such, Freeman offers a compositional scheme that allows for representing the structure of argument *presentations* in text. The main devices (which prove sufficient for handling a dozen of texts from the PCC, which so far have been analyzed) are

- *text units*, indicated by thick horizontal lines,
- *support* relations between units, indicated by lines and arrows,
- *dependence* between units that only collectively support another unit (rather than individually), indicated by a horizontal line connecting the units,
- *counter-arguments* that block a support relation, indicated by thicker lines drawn through the support-line, and
- *refutations* of counter-arguments, drawn here as arcs and arrows.

Consider the analysis of the sample text, given in Fig. 6. It identifies (21) as the central claim of the author, and a series of supporters that also recursively support one another: The successful fight against pox (8–9) directly supports the pro-vaccination statement, as does the achieved reduction of other infections (14–15). The latter is in turn supported by the stated contrast between death tolls in developing countries (11) and the infrequence of infections in Germany (12–13). Since these need to be considered as contrasting to make the argument work, they are an instance of *dependent* units. More support comes from the appeal that our children should not suffer needlessly (19–20), which is in turn supported by (17–18). The counterarguments and their refutations have been mentioned above; in this particular case they counter not one particular support relationship but run against the author's thesis in general; therefore, the "blocked" arrow [the rightmost one leading into (21)] does not originate in any specific supporting material.

4 Corpus Annotation with MLA

Having argued that MLA is an appropriate discourse representation from a descriptive viewpoint, we now turn to empirical analysis and the task of corpus creation. This

involves the technical question of software infrastructure on the one hand, and the design of an incremental annotation procedure on the other.

4.1 Technical Infrastructure

As long as a corpus involves a single level of analysis, data can be created, and often also viewed, with a single tool. The MLA approach, on the other hand, can only be realized effectively when dedicated task-specific annotation tools are employed for the various levels, and when the set of all resulting annotation layers can be merged into a single data structure for visualization, querying, and statistical analysis.

As for annotation, we are using several existing tools: *annotate* (Brants and Plaehn 2000) for (semi-automatically) annotating syntactic structure; *Exmaralda* (Schmidt 2004) for minimal units, illocutions and thematic structure; MMAX2 (Müller and Strube 2006) for coreference; *ConAno* (Stede and Heintze 2004) for conjunctive relations; and *RSTTool* (O'Donnell 2000) for intentional structure. All these tools produce their own XML output formats, which we map to a generic standoff annotation representation (*PAULA*; Dipper 2005) by conversion scripts. The linguistic information system *ANNIS* reads the standoff annotation layers, stores them as Java objects in main memory (and in future also persistently in an underlying relational database), provides visualization of the individual levels upon request, and offers a query language that allows for cross-level searches. For a detailed description of this framework, see Chiarcos et al. (2008).[5]

An important feature of this architecture is its being open for additional annotation layers—they can be added anytime by conversion to PAULA and incremental import into ANNIS. In this way, information that has not been annotated at first but later turns out to be useful can be added and set into correspondence with the existing annotations for the text.

4.2 Annotation Practice for the Potsdam Commentary Corpus

We have annotated a portion of the (German) PCC according to MLA, and a set of sample texts in English is currently in preparation. In the following, we briefly describe the annotation process.

After the preprocessing step of automatic tokenization (as the basis for all annotations and their subsequent merging), sentence syntax is annotated according to the TIGER scheme (Brants et al. 2004), which is a relatively theory-neutral approach to syntactic structure, and which has been successfully applied to large corpora of newspaper texts. This step is performed by experienced annotators who have had extensive training in syntactic issues. The output of the analysis tool is mapped to ANNIS, but also filtered for noun phrases, which serve as potential markables for coreference annotation.

[5] Both the conversion scripts and the database tool are being made available to interested researchers. Please contact the author.

All annotation levels other than syntax are then produced by a single annotator, who addresses the task in a fixed order. Coreference is annotated according to the 'PoCoS' guidelines developed by Krasavina and Chiarcos (2007) for German, which in turn build upon the MATE/GNOME scheme for English (Poesio 2004). Krasavina and Chiarcos suggest a two-tier appoach: In the "basic" scheme, only direct nominal anaphora is handled, and in the "extended" scheme, indirect or 'bridging' anaphora is also accounted for. Event anaphora, however, is excluded for the time being. The annotator then turns to thematic structure; our guidelines ask to provide a spanning analysis that includes every portion of the text as belonging to some thematic unit. Using the tier-oriented Exmaralda tool, the annotator can easily mark formulaic segments and topic-introducing clauses, and define thematic units, possibly with embedding. Next, the results are added to the text representation in ANNIS, and the annotator then uses Exmaralda to perform segmentation into minimal units and assigning illocutions to them. For segmentation, we use the guidelines by Jasinskaja et al. (2007) for German, which took some inspiration from those by Carlson and Marcu (2001) for English.

For conjunctive relations, the ConAno tool aids the annotator by automatically suggesting connectives to be annotated (see Stede and Heintze 2004). As for the spans, they are marked by the annotator in the raw text; we do not prescribe syntactic units as possible spans. Our guidelines are essentially a compact version of the description by Martin (1992). The analyst works through the text left-to-right and considers just one segment relationship at a time—she does not have a complete view of all relation assignments. This is different with the argument structure, where it is obviously important to consider the role of a segment with respect to the entire argumentation. Thus, a tool is used that allows for incrementally building up the tree.

At any time in the annotation process, the annotator can use ANNIS to inspect the results already produced on "lower" levels. Thus in the end, we regard the set of discourse-level annotations as the complete view that an individual annotator has formed of the text.

5 Related Approaches

In this paper, we have taken Rhetorical Structure Theory as the starting point, because it is still quite influential in empirical discourse research. Nonetheless, some alternative approaches have been mentioned in passing already; in this section, we briefly summarize their relationship to our proposal of MLA. In general, the attitude towards "deconstructing" complex phenomena into a series of smaller, less complex phenomena is shared, for different aspects, by various authors, for example: Bateman (2001) discusses the relationship between rhetorical structure and thematic structure; he also sees conjunctive relations as a surface-oriented counterpart of some "deeper" relational representation. Knott et al. (2001) advocate removing the *Elaboration* relation from RST trees and instead add a device called 'entity chains', which seems to be especially useful for the text genre they consider (descriptions of museum exhibits). A similar proposal is made by Poesio et al. (2004), who furthermore suggest to decompose the annotation of Centering theory transitions into simpler subtasks.

As for accounts of "relational" discourse coherence that have lead to annotated corpora (besides the aforementioned RST Treebank), the 'Penn Discourse Treebank' (PDTB; Prasad et al. 2004) employs a scheme of annotating connectives that is quite similar to that of Martin (1992), though his work is not explicitly referred to. Thus the PDTB approach is also very similar to the conjunctive relations level of MLA, with our point being that explaining coherence requires additional levels. The 'Discourse Graph Bank' (Wolf and Gibson 2005) also criticizes RST for its relying on a single tree and instead advocates more general graph structures; the authors stress in particular the formal question of the descriptive power of trees versus graphs. In their representations, more links are encoded than in RST, and they are all assembled in a graph structure. In comparison to MLA, the systematic division into conceptual realms is not accounted for. Finally, 'Relational Discourse Analysis' (RDA; Moser and Moore 1995) realizes a two-level approach by distinguishing informational and intentional relations, somewhat tailored to the targeted domain. But RDA explicitly affirms nuclearity as organizing principle for both levels. While this might be useful for their text genre (tutorial dialogue), we had already pointed to Stede (2008) for a critical discussion of the general role of nuclearity. Besides, the informational relations do not quite correspond to conjunctive relations, as they are meant to capture domain semantics, which the CRs do only to some extent. Also, RDA does not account for explicit illocutions nor for the roles of referential and thematic structure.

6 Summary

Coherence is a complex phenomenon: A text is *globally* coherent when it is perceived as a thematic whole and fulfils a recognizable overall function; it is *locally* coherent when it is perceived as connecting the units of information in a purposeful and sensible way, and when a felicitous balance between continuity and newness is achieved in the linear unfolding. And, of course, a text is not an isolated object: For perceiving it as coherent, the reader brings to bear his or her linguistic knowledge as well as world and domain knowledge. In order to better understand this complexity, it is advisable to follow a divide-and-conquer strategy by isolating sub-phenomena and constructing models for them. The crucial first question, however, is how to divide. The proposal of MLA is to cut "horizontally": to represent the text as a stack of layers representing different realms of information, some of which rather directly reflect linguistic features, and some of which involve subjective interpretation.

An RST-style analysis of a text, on the other hand, cuts "vertically": It tries to capture the essence of coherence within a single representation structure, making a series of quite different simplifications along the way. We do not doubt that this can be an insightful instrument for studying text—RST has been quite successful for a variety of purposes. But there are inherent limitations on the explanatory power when information from different realms is conflated in a single tree structure: On the one hand, one cannot do full justice to the separate realms; on the other hand, the single tree structure becomes ambiguous, because when crafting it, many underlying assumptions cannot be made explicit.

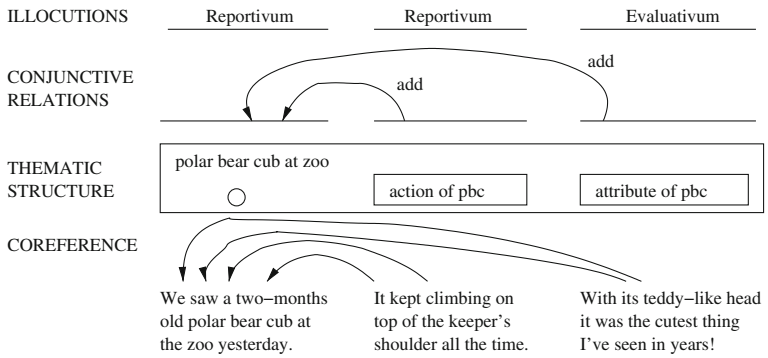

Fig. 7 MLA representation of example (1)

For a final illustration, Fig. 7 shows the MLA account of example 1. The referential structure is again simplified, it does not identify the precise coreferring expressions. The other levels all capture various aspects of the alternative RST trees that we had listed in Sect. 1. Thematic structure notes the topic-setting function of the first sentence and the subtopics of the subsequent sentences, which corresponds to the information expressed by a *Topic-Comment* relation in RST (but explicitly adds the information about sub-topics). Conjunctive relations indicate the relatively loose overall structure—two separate *Additions* mirror an RST tree with two distinct *Elaboration* relations. The assignment of illocutions captures the evaluation in the third sentence. Finally, the fact that our alternative RST analyses display a great variety of nuclei variation points to the fact that there really is *no* nuclearity involved in the text. The only "imbalance" is the one captured on the levels of thematic structure and conjunctive relations. Proper nuclearity in MLA (support relations between speech acts) is absent, and hence the intentional structure for the text is empty. This means that each segment plays its illocutionary role by itself; connections are perceived in terms of thematic development, but not in terms of intentions. Nonetheless, the text is perfectly coherent, which illustrates that a complex structure on the intentional level is indeed not a necessary condition for a "good" text.

Studying coherence can greatly benefit from annotated corpora, and we have argued that a systematic multi-level annotation is of greater utility than one where analysts have to annotate a complex phenomenon in one shot. Given a dedicated software environment, the "simpler" levels of annotated text can be combined to look into more difficult levels: One avenue is to investigate correlations between coreference and thematic structure, which in turn can be linked to syntactic structure and shed more light on constraints and preferences for arranging the information structure of sentences in an unfolding text. Or, to mention just one more application, the expression of subjectivity and opinion in text can be studied as an interplay of syntactic constructions and intentional relations, plus additional information on subjective lexical elements, which would form another level of annotation.

Hence, the four levels we have described here are of course not the only interesting and useful ones; but our claim is that they indeed play central roles for text coherence, and that the information conflated in an RST-like tree can be effectively decomposed

into these four realms. Empirical studies on annotation efficiency and inter-annotator agreement, based on level-specific guidelines, are currently in preparation.

Acknowledgements The author is grateful to two anonymous reviewers and to the guest editors for their constructive suggestions on improving earlier versions of this paper.

References

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.

Bateman, J. (2001). Between the leaves of rhetorical structure: static and dynamic aspects of discourse organization. *Verbum, 23*(1), 31–58.

Bateman, J., & Rondhuis, K. J. (1997). "Coherence relations": Towards a general specification. *Discourse Processes, 24*(1), 3–50.

Brandt, M., & Rosengren, I. (1992). Zur Illokutionsstruktur von Texten. *Zeitschrift fr Literaturwissenschaft und Linguistik, 86*, 9–51.

Brants, S., Dipper, S., Eisenberg, P., Hansen, S., König, E., Lezius, W., Rohrer, C., Smith, G., & Uszkoreit, H. (2004). TIGER: Linguistic interpretation of a german corpus. *Research on Language and Computation, 2*(4), 597–620.

Brants, T., & Plaehn, O. (2000). Interactive corpus annotation. In *Proceedings of Second International Conference on Language Resources and Evaluation (LREC-2000)*. Athens.

Carlson, L., & Marcu, D. (2001). Discourse tagging reference manual. Technical report, University of Southern California/ISI.

Carlson, L., Marcu, D., & Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In J. van Kuppevelt & R. Smith (Eds.), *Current directions in discourse and dialogue*. Dordrecht: Kluwer.

Chiarcos, C., Dipper, S., Götze, M., Ritz, J., & Stede, M. (2008). A flexible framework for integrating annotations from different tools and tagsets. In *Proceedings of the First International Conference on Global Interoperability for Language Resources*. Hongkong.

Dipper, S. (2005). XML-based stand-off representation and exploitation of multi-level linguistic annotation. In R. Eckstein & R. Tolksdorf (Eds.), *Proceedings of Berliner XML Tage* (pp. 39–50). Berlin: Humboldt University.

Figge, U. L. (1971). Syntagmatik, Distribution und Text. In W.-D. Stempel (Ed.), *Beiträge zur Textlinguistik* (pp. 161–181). München.

Freeman, J. B. (1991). *Dialectics and the macrostructure of argument*. Berlin: Foris.

Grosz, B., & Sidner, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12*(3), 175–204.

Jasinskaja, K., Mayer, J., Boethke, J., Neumann, A., Peldszus, A., & Rodríguez, K. J. (2007). Discourse Tagging Guidelines for German Radio News and Newspaper Commentaries. Ms., Universität Potsdam.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Knott, A., Oberlander, J., O'Donnell, M., & Mellish, C. (2001). Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 181–196). Amsterdam: John Benjamins.

Krasavina, O., & Chiarcos, C. (2007). PoCoS: The Potsdam Coreference Scheme. In *Proceedings of the Linguistic Annotation Workshop (LAW) at ACL-07*. Prague.

Lötscher, A. (1987). *Text und Thema. Studien zur thematischen Konstituenz von Texten*. Tübingen: Niemeyer.

Mann, W., & Thompson, S. (1988). Rhetorical structure theory: Towards a functional theory of text organization. *TEXT, 8*, 243–281.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. Cambridge/MA: MIT Press.

Martin, J. R. (1992). *English text: System and structure*. Philadelphia/Amsterdam: John Benjamins.

Matthiessen, C., & Thompson, S. (1988). The structure of discourse and 'subordination'. In J. Haiman & S. Thompson (Eds.), *Clause combining in grammar and discourse* (pp. 275–329). Amsterdam: John Benjamins.

Moore, J., & Pollack, M. (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics, 18*(4), 537–544.

Moser, M., & Moore, J. (1995). Using discourse analysis and automatic text generation to study discourse cue usage. In *AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, pp. 92–98.

Müller, C., & Strube, M. (2006) Multi-level annotation of linguistic data with MMAX2. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy* (pp. 197–214). Frankfurt: Peter Lang.

O'Donnell, M. (2000). RSTTool 2.4—a markup tool for rhetorical structure theory. In *Proceedings of the International Natural Language Generation Conference* (pp. 253–256). Mizpe Ramon/Israel.

Poesio, M. (2004). The MATE/GNOME proposals for anaphoric annotation, revisited. In *Proceedings of the SIGDIAL '04 Workshop*. Cambridge/MA: Association for Computational Linguistics.

Poesio, M., Stevenson, R., di Eugenio, B., & Hitzeman, J. (2004) Centering: A parametric theory and its instantiations. *Computational Linguistics, 30*(3), 309–363.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics, 12*, 601–638.

Prasad, R., Miltsakaki, E., Joshi, A., & Webber, B. (2004). Annotation and data mining of the Penn discourse Treebank. In *Proceedings of the ACL Workshop on Discourse Annotation*. Barcelona.

Redeker, G., & Egg, M. (2006). Says who? On the treatment of speech attributions in discourse structure. In *Proceedings of the Workshop on Constraints in Discourse*. Maynooth University/Ireland.

Sanders, T., & Spooren, W. (1999). Communicative intentions and coherence relations. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Coherence in Spoken and Written Discourse*. Amsterdam: Benjamins.

Sanders, T., Spooren, W., & Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes, 15*, 1–35.

Schmidt, T. (2004). EXMARaLDA-ein System zur computergestützten Diskurstranskription. In A. Mehler & H. Lobin (Eds.), *Automatische Textanalyse* (pp. 203–218). Wiesbaden: Verlag für Sozialwissenschaften.

Schmitt, H. (2000). *Zur Illokutionsanalyse monologischer Texte*. Frankfurt: Peter Lang.

Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society, 5*, 1–23.

Stede, M. (2004). Kontrast im Diskurs. In H. Blühdorn, E. Breindl, & H. Wassner (Eds.), *Brücken schlagen. Grundlagen der Konnektorensemantik* (pp. 255–286). Berlin: Walter de Gruyter.

Stede, M. (2008). RST revisited: Disentangling nuclearity. In C. Fabricius-Hansen & W. Ramm (Eds.), *'Subordination' versus 'coordination' in sentence and text*. Amsterdam: John Benjamins.

Stede, M., & Heintze, S. (2004). Machine-assisted rhetorical structure annotation. In *Proceedings of the 20th International Conference on Computational Linguistics* (pp. 425–431). Geneva.

Taboada, M., & Mann, W. (2006). Rhetorical structure theory: Looking back and moving ahead. *Discourse Studies, 8*(4), 423–459.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

van Dijk, T. (1977). *Text and context. Explorations in the semantics and pragmatics of discourse*. London, NY: Klett.

Webber, B., Stone, M., Joshi, A., & Knott, A. (2003). Anaphora and discourse Structure. *Computational Linguistics, 29*(4), 545–587.

Wolf, F., & Gibson, E. (2005). Representing discourse coherence: A corpus-based study. *Computational Linguistics, 31*(2), 249–287.