

*Rhetorical Structure Theory:
A Review*

MAITE TABOADA
SIMON FRASER UNIVERSITY, CANADA
WILLIAM C. MANN¹
INDEPENDENT RESEARCHER

ABSTRACT Rhetorical Structure Theory has enjoyed continuous attention

Citations are in a wide range of disciplines, from teaching English as a second language (Connor, 2002) and applied linguistics (O'Brien, 1995) to knowledge management (Gil and Ratnakar, 2002) and crisis negotiation (Taylor, 2002), through artificial intelligence and computational linguistics (e.g. Afantenos et al., 2004; Grasso, 2003; Hovy, 1988; Wahlster et al., 1991), among many others.

It was with this popularity in mind that we decided to do a review of some of the praise and criticism that RST has received over the years, also paying attention to extensions, modifications and applications. A few summaries exist, but they often focus on one particular application or field of study (Károly, 1998; Thomas, 1995), or are meant for a more general audience (Bateman and Delin, 2005). It was, however, soon evident that a completely exhaustive review of all applications and extensions was impossible. We decided to concentrate on a few

developed without strong links to any previous descriptive tradition. It was intended for a particular kind of use, to guide computational text generation, but that use did not strongly influence the framework. As noted below, it has been used in very different ways (see also Taboada and Mann, forthcoming).

The theory started with few assumptions about how written text functions, and how it involves words, phrases, grammatical structure, or other linguistic entities (Mann et al., 1992; Matthiessen and Thompson, 1988). This agnostic beginning was crucial in shaping the result. RST is intended to complement other text description methods. The most familiar kinds of linguistic description, about words, phrases, grammatical structure, semantics and pragmatics, all make contributions that are qualitatively distinct from those of RST. The theory was defined in a flexible, open way as a tool that could be adapted to various applications and linguistic situations.

The introduction to RST below is quite simplified. For a more extensive introduction, see the RST website (Mann, 2005), or the published description of RST (Mann and Thompson, 1988). The website also contains a scan of a much more complete technical report (Mann and Thompson, 1987) that was the basis for the published article.

2.1 THE MEANING OF RELATIONS

RST addresses text organization by means of relations that hold between parts of a text. It explains coherence by postulating a hierarchical, connected structure of texts, in which every part of a text has a role, a function to play, with respect to other parts in the text. The notion of text coherence through text relations is widely accepted, and the relations have also been called *text relations*, *textual relations*, or *textual coherence relations* in the literature. Asher and Lascarides (2003) use the term *textual coherence relations*, although their theory is different from RST.

RST provides a systematic way for an analyst (also called observer or judge) to annotate a text. If the annotation involves an entire text, or a fairly independent fragment, then the analyst seeks to find an annotation that includes every part of the text in one connected whole. An analysis is usually built by reading the text and constructing a diagram that resembles Figure 1. This is a title and summary, appearing at the top of an article in *Scientific American* magazine (Ramachandran and Anstis, 1986). The original text, broken into numbered units, is:

1. [Title:] The Perception of Apparent Motion
2. [Abstract:] When the motion of an intermittently seen object is ambiguous,
3. the visual system resolves confusion
4. by applying some tricks that reflect a built-in knowledge of properties of the physical world.

The main way in which one unit becomes connected to another is by adding an RST relation to the diagram. An example in Figure 1 is represented by the

arrow from unit 2 to a span labelled 3–4. A span of text is at each end of an arrow, and the arrow is labelled with the name of a relation, in this case Condition. The arrowhead points to a span called the nucleus (units 3–4) and the arrow points away from another span called the satellite (unit 2). All units are also spans, and spans may be composed of more than one unit. The abstract representations in this type of diagrams are referred to as schemas.

RST relations are defined in terms of four fields: 1) Constraints on the nucleus; 2) Constraints on the satellite; 3) Constraints on the combination of nucleus and satellite; and 4) Effect (achieved on the text receiver). To specify each field for any instance of a particular relation, the analyst must make plausibility judgments, based on context and the intentions of the writer. All relations are defined in terms of the four fields. Definitions are based on functional and semantic criteria, not on morphological or syntactic signals, because no reliable or unambiguous signal for any of the relations was found.

Different lists of relations exist. The original set defined in the 1988 article includes 24 relations,⁴ dubbed ‘Classical RST’ (Nicholas, 1994; Rösner and Stede, 1992). More recent work has added definitions for List, Means, Preparation, Unconditional, and Unless. Also, it has given Restatement both nuclear and multinuclear forms, reflecting more recent understandings of particular texts. This, along with Joint (the declared absence of a relation) yields a total of 30 (Mann, 2005).

2.2 NUCLEARITY IN DISCOURSE

RST establishes two different types of units. Nuclei are considered as the most important parts of a text, whereas satellites contribute to the nuclei and are

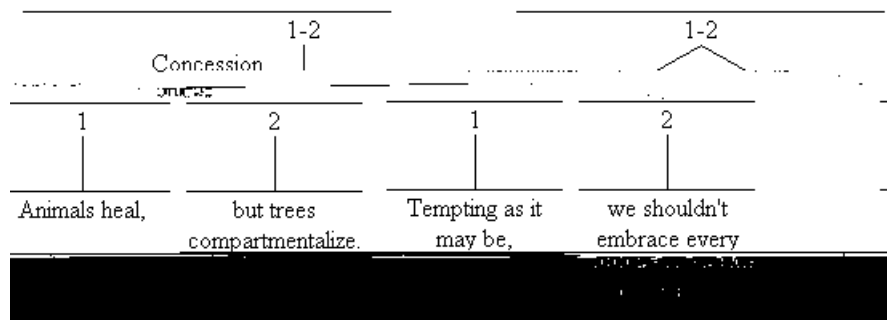


FIGURE 2. Concession and Contrast

be the basic information, and the satellite contains additional information about the nucleus. The nucleus is more essential to the writer's purpose than the satellite. The satellite is often incomprehensible without the nucleus, whereas a text where the satellites have been deleted can be understood to a certain extent.

This nuclearity principle is also the basis of hypotactic relations postulated for lower levels of organization in language (i.e. the main-subordinate distinction in complex clauses). Languages also exhibit parataxis, the coordination of structures of equal importance. In RST, parataxis is reflected in multinuclear relations, those where no span seems more central than the other to the author's purposes. In fact, Matthiessen and Thompson (1988) argue that what is commonly referred to as subordination at the clause level may have arisen out of the grammaticalization of rhetorical relations.

Figure 2 presents two examples, one of a nucleus-satellite relation, Concession, and the second of a multinuclear relation, Contrast, both taken from the RST website. In the Concession relation, the nucleus ('we shouldn't embrace every popular issue that comes along') is considered to be more central. On the other hand, Contrast joins two units that seem to be of equal importance.

2.3 HIERARCHY IN THE ANALYSIS

The hierarchy principle in RST is part of the nucleus-satellite distinction described in the previous section. RST relations are applied recursively to a text, until all units in that text are constituents in an RST relation. This is because the effect to be achieved with a particular relation may need to be expressed in a complex unit that includes other relations. The effect of one particular text can be summarized in one top-level relation, but decomposed in further relations that contribute to that effect.

Analysis of a text is performed by applying schemas that obey constraints of completeness (one schema application contains the entire text); connectedness (each span, except for the span that contains the entire text, is either a minimal unit or a constituent of another schema application); uniqueness (each schema application contains a different set of text spans); and adjacency (the spans of each schema application constitute one contiguous text span).

The result of such analyses is that RST structures are typically represented as trees, with one top-level relation that encompasses other relations at lower levels, as illustrated in Figure 1, where the main relation is one of Preparation (the title

generation of texts. It was also meant to be a general theory of how text works, and how coherence in text is achieved. Over the years, RST has been adopted by many researchers for very different purposes, making it difficult, at this point, to describe RST in terms of one or two main goals. What we can propose is a set of characteristics that a theory such as RST would ideally possess, leading to explanations for the roles that RST can play. In initial descriptions of RST, and in subsequent discussions, two main characteristics are proposed: descriptive

becomes a source of problems if the text is multilingual, or if it is in a language that is not as clause-centred as English (i.e. where clausal units are not easily established).

- (1) My flicking the switch caused the light to turn on.

Nicholas (1994) discusses the issue at length, addressing not only nominalizations, as in the example, but also complement clauses and relative clauses. This downward granularity may be desirable in some applications, such as Natural Language Generation (Vander Linden et al., 1992).

3.3 COHERENCE, HIERARCHY, INTENTIONS, AND THE NATURE OF RELATIONS

The RST definition of coherence of text involves finding an intended role in the text for every unit. Negatively, coherence is the absence of non-sequiturs. In order to find roles for every part of each text, and noting the very local character of most of these roles, recognition of hierarchy seems essential. In identifying such hierarchy, RST combines the idea of nuclearity (higher relative importance) with the identification of relations. Experience shows this to be a questionable combination. For relations like Background, the idea that the background information is never the most important is quite credible. For others, such as Volitional Cause (or Volitional Result), disassociation of nuclearity from the relation identity seems to be a useful step. An alternative version of RST done in this manner would be interesting.

Hierarchy and nuclearity have been most important in computational applications of RST. Marcu (1996, 1997) proposed a compositionality principle: 'whenever two large text spans are connected through a rhetorical relation, that rhetorical relation holds also between the most important parts of the constituent spans' (i.e. the nuclei and their daughter relations, recursively). Inversely, when building up rhetorical structure trees for text (e.g. in parsing text), a discourse relation between two large text spans can only be postulated if the relation also holds between the most important units in the spans. This compositionality principle has been criticized as unmotivated (Wolf and Gibson, 2004b), and as insufficient for planning text structures (Bouayad-Agha, 2000), or for summarization, even by Marcu himself (Marcu, 1998b).

If we assume, however, the notions of hierarchy and compositionality for now (see Section 3.4 for further discussion), the next question is what is that is organized in hierarchical manner to produce coherence: do the texts spans represent propositions, intentions, or something else? Research in discourse, especially in computational applications, has focused on the notion of intentionality as a source of coherence. Grosz and Sidner (1986) propose intentions as one of the underlying components of discourse. The study of intentions in discourse spawned much research in the 1990s, some of it related to the general problem of planning in Artificial Intelligence (Bratman, 1987; Cohen et al., 1990; Grosz and Sidner, 1990; Litman and Allen, 1990), and in particular to the planning of natural text. A 1993 workshop sponsored by the Association for Computational Linguistics addressed the issue of 'Intentionality and Structure in Discourse Relations' (Rambow, 1993). In the introduction to the proceedings, Rambow points out that RST relations may be simply a reflection of intentions

and relations among intentions. We believe that intentions and relations are different, although possibly complementary. Rhetorical relations join two spans in order to create a new span. The reason for connecting the spans is to create an effect on the reader. That effect may well be considered an intention. Intentions, however, can also be satisfied by uttering a single sentence, or even a single word. Sanders and Spooren (1999) argue that intentions are poorly defined, and are not on a par with coherence relations. Intentions are unary functions that apply to a single proposition. Relations, on the other hand, apply to at least two propositions (or to two 'chunks' that may include a number of propositions). Intentions, by themselves, cannot account for the coherence of discourse. Asher and Lascarides (1994) also separate intentions from discourse structure and

section, however, we would like to discuss a number of studies that have shown the importance of relations in the processing of discourse.

Sanders et al. (1992) carried out psycholinguistic experiments in order to test a taxonomy of coherence relations (see Section 3.5). Although their goal was not to test psychological validity, from the experiments it can be gathered that subjects were sensitive to different relations. They also report on other experiments that suggest that coherence relations, especially when explicitly marked, help organize discourse representation, as evidenced in off-line reproductions (Meyer et al., 1980); and that linguistic markers of coherence

faster

immediately denies the Relational Proposition, then the text may appear incoherent. In Example (2), the apparent relation has to do with cause and therefore the Relational Proposition asserts cause. If the text continues with another sentence, as in (3), the assertion of cause is denied and the corresponding causal relation cannot be part of the analysis. If there is no other plausible way to relate the sentences, the text will seem incoherent.

(2) The bottle fell. It broke.

(3) The bottle fell. It broke. I don't know why it broke.

Coherence, in this view, stems from the relation between two (or more) chunks of discourse. This idea has been referred to as *entity-based coherence*, as opposed to *relation-based coherence*. The principle is that coherence is created through two different, but related mechanisms: the presence of entities that form chains in a discourse (entity-based coherence), and the presence of implicit or explicit relations between the parts that form a text. The dual mechanism can be traced back to at least Halliday and Hasan's (1976) notion of cohesion, which includes not only non-structural relations in a text (repetition, substitution, synonymy, etc.), but also conjunctive relations, in their case always signalled by a marker (additive, adversative, causal, temporal). The relationship between reference and structure has been explored before (e.g. Fox, 1987), but it is only recently that the two have been considered as aspects of general coherence (Kehler, 2002; Poesio et al., 2004). RST would, then, be only one part of that whole (Kibble and Power, 2004).

There is a line of research exploring the relationships between anaphora, discourse structure and syntax. Identifying discourse relations and resolving anaphoric reference represent two kinds of text interpretation problems with comparable, often multisentential, scales of action. The results of these two kinds of processes must be consistent, and so the processes potentially can help each other by supplying constraints that accompany various alternatives of interpretation (Cristea et al., 2000; Kehler, 2002; Kruijff-Korbayová and Webber, 2000; Schauer and Hahn, 2001; Webber et al., 1999; Wolf et al., 2004). A more radical approach (Webber et al., 2003) suggests that relations introduced by certain discourse adverbs are not structural at all, but only anaphoric. Webber and colleagues propose a new theory of anaphora in which such adverbials can be considered as anaphors.

We hope that further studies on the relationship between anaphora and coherence, and on coherence in general, will help shed light on the issue.

3.4 SHOULD RST ANALYSES BE RESTRICTED TO TREES?

RST analyses, including the diagrams in Figures 1 and 2, are generally presented as trees. Some authors have said that trees are the only structures allowed in RST.⁶ Even the RST Tool⁷ produces only tree diagrams. In fact, much of the computationally-oriented research on discourse postulates tree structures

(Asher and Lascarides, 2003; Grosz and Sidner, 1986; Polanyi, 1988; Webber, 2004). For many researchers, this seems to be a flaw.

In fact, the set of RST-defined relations in particular texts often cannot be represented by single trees. Parallelism is one of the earliest recognized kinds of cases. In Mann and Thompson (1988: 265), one of the cited varieties of multiple analyses in RST was 'simultaneous compatible analyses'. These arise when one analyst accepts multiple analyses and says, in effect, that the claims in each of the analyses hold. This is enough machinery to represent parallelism in texts, provided only that the relations that form the parallelism are made to be RST relations. The resulting structure cannot be a tree any longer. This approach also handles the case of multifunctional relations (see Section 3.7).

Chafe (1996) discusses the problems of using trees to represent discourse, particularly spoken discourse:

A tree diagram falls short of capturing the gradual development of ideas through time under the influence of both cognitive and social goals and constraints. People move from one thought to another, chunking certain thoughts together to be sure, but continually influenced by ongoing processes of memory as well as by the

Restatement) and presentational relations (Motivation, Background, Justify, Concession). The classification is based on the effect intended: in subject matter relations the text producer intends for the reader to recognize the relation; in presentational relations the intended effect is to increase some inclination on the part of the reader (positive regard, belief, or acceptance of the nucleus). The distinction is related to the semantic/pragmatic divide proposed by van Dijk to classify discourse connectives according to what type of relation they signal: '[p]ragmatic connectives express relations between speech acts, whereas semantic connectives express relations between denoted facts' (van Dijk, 1979: 449). The distinction has received numerous labels:

- External/internal (Halliday and Hasan, 1976; Martin, 1992).
-

Taboada, 2004a, in press) have found that the following relations are never marked: Enablement, Evaluation, Elaboration and Solutionhood. Rarely marked are: Background and Summary.

Mann and Thompson (1988) proposed other classifications, based on where the locus of effect is (nucleus or satellite), whether the relation involves reader action, or the most common order of constituents (nucleus first or satellite first). Further classifications are also possible.

As for the number of relations, the original list of 24 relations, expanded to 30 on the RST website (Mann, 2005) could have been shorter, reflecting hierarchical structure without reflecting different functions for different parts of the structure. For various reasons, the list could not have been radically longer and still effective. Perhaps the strongest limit on distinguishing various relations in analysis is the possibility of distinguishing one relation from another, sometimes called

A particular problem area for this is the Elaboration relation, to the point that there is a proposal to remove it from the set of rhetorical relations (Knott et al., 2001). It has six conditions in the definition, any one of which allows someone to find that Elaboration is present (e.g. set::member and abstraction::instance). In principle, if all cases were clear, there could be six relations instead of one. Marcu and colleagues (Carlson and Marcu, 2001) followed precisely that approach. But it is very frequent that the finding of Elaboration is clear, and the identification of the subtype is not. Forcing a decision is not useful. So, for analysis, there is one relation, and the distinctions are lost during analysis.

There have been a number of efforts to develop alternate collections of relations, distinct from the RST relations and founded on some alternate basis. For instance, Hovy and Maier (Hovy and Maier, 1995; Maier and Hovy, 1991, 1993) propose to use language metafunctions (ideational, interpersonal and textual), to arrive at a total of about 70 relations, applied to text generation (Hovy et al., 1992). A similar approach is the use of Halliday's (1985) classification of relations among clauses into elaboration, enhancement and extension (Maier and Hovy, 1993). This work is based on a large collection of about 350 relations from various sources (Hovy, 1990). Grosz and Sidner (1986) propose that two relations (dominance and satisfaction-precedence) are sufficient to characterize discourse structure. For a summary of these and other proposals, see Nicholas (1994).

Sanders and his colleagues (Sanders et al., 1992, 1993) have mapped the space of possible relations to a number of primitive concepts. The primitive

classification, Sanders et al. asked subjects to label texts (in Dutch) with the appropriate relation. The results show moderate agreement in both cases, measured according to the kappa coefficient (Carletta, 1996; Hubert, 1977). The subjects tended to agree least on the source of coherence, and most on the polarity. In terms of classes, there is frequent disagreement between contrast and concession, condition–consequence and goal–instrument.

Sanders et al. (1992) propose a number of areas in which the taxonomy can be further investigated. The first one is language acquisition. Some studies suggest that children acquire relations and their connectives in a fixed order (e.g. Bloom et al., 1980). If that is the case, it would point to classes of relations, as Sanders and colleagues suggest. A second area of research is text analysis. Extensive analyses would show the suitability of possible taxonomies. Unfortunately, in the years since this proposal, we know of no extensive corpus analysis to test a taxonomy. A number of studies have used RST or related theories to perform text analysis, but they have relied on existing taxonomies (Taboada, 2004b), or they have created their own, but not tested them against other possibilities (Marcu, 2000a, 2000b).

The various efforts to taxonomize relations, to find or introduce hierarchy among RST relations, and to classify relations into types, are collectively supportive 9.6(i tyw)38.6(e)0(k4 9.N0as or fulnondhe)JTJT*-tie.7(s pg 9.N0a.7(all-f 98.8(pprIt.7(mm)-3 1461

argued (Scott and de Souza, 1990) that all rhetorical relations proposed in the 1988 article can be signalled, whether through lexical or phrasal units or through syntactic relations (e.g. embedding). The claim was made of two languages: British English and Brazilian Portuguese, and it could possibly be extended to other languages. In this section, we discuss studies of relation signalling, and what the generalized absence of signalling means for RST. We will not attempt to review all the research on coherence relation signalling, since the studies are too numerous to be mentioned here. A few cases will suffice to exemplify the depth and diversity of the research.

Cue phrases have been the main object of study in the area of relation signalling. They have received different names: coherence markers, cue phrases, discourse connectives, or discourse markers. For examples and definitions, see, among others, Blakemore (1987, 1992, 2002), Fraser (1999), or Louwerse and Mitchell (2003), Redeker (1990, 1991), or Schiffrin (1987, 2001). The study of discourse markers is an extensive area of research in itself. It has been characterized as 'a growth industry in linguistics' (Fraser, 1999: 932).

Knott and colleagues have produced a series of studies in which they motivate a taxonomy of coherence relations based on cue phrases, including a study of the taxonomy across languages, English and Dutch (Knott, 1996; Knott and Dale, 1994, 1996; Knott and Mellish, 1996). Their rigorous methodology has as a starting point the collection of a large corpus of candidate connectives (a total of about 150), and continues with a classification of the connectives, by means of substitution tests. Although the classification is described as preliminary, the taxonomy is already quite complex and sophisticated. One possible next step in this line of research is to use the taxonomy of connectives to create a taxonomy of relations. Sanders and colleagues (1992) are also interested in the adequacy of a taxonomy, and in the psychological plausibility of coherence relations.

Taboada (2001, 2004a) compared conversations in English and Spanish, and their signalling, and found that relations signalled by conjunctions (coordinating and subordinating) were always local relations at a lower level in the discourse, with no signalling for relations that spanned across turns. It is likely that signals such as conjunctions occur at a more local level, where they may have become grammaticalized (Matthiessen and Thompson, 1988). Another study (Taboada, in press), of *Summary* articles from the RST corpus (Carlson et al., 2002), shows levels of signalling ranging from 4 percent of all Summary relations (a total of 75) to over 90 percent in 228 Concession relations.

Schauer and Hahn (2001) discuss the manual coding of a German corpus with a total of 549 relations. Of those, about 61 percent were unmarked. They conclude that cue phrases, by themselves, are not good indicators of coherence relations in text, and explore anaphoric relations as indicators of rhetorical structure. They argue that two units that contain co-referential relations have to be related in a way that helps solve the reference. This assumption is the basis of an algorithm that helps construct discourse structure trees. When processing a new unit, the unit is attached to the previous unit that contains some form of co-

reference to the unit being processed. Although cue phrases are also used, the addition of co-reference results in discourse structure correctly predicted in 64 percent of the cases (no baseline comparison is given). The authors also mention that excluding confounding cases of intra-sentential cue phrases could result in a correct discourse structure for 86 percent of the cases. One question that arises here is whether we can consider co-reference as an indicator of rhetorical relations, or as a parallel system. In other words, it is a question of whether cohesion and coherence are parallel systems or whether they are interrelated, one supporting the other.⁹

The question we are interested in here is the issue of unsignalled relations. One position is that the fact that unsignalled relations exist does not mean that the relations are not present, in the same way that zero anaphora does not mean that an anaphoric relation is not present. The issue is, then, how to classify those relations that are not overtly signalled. Knott and Dale (1994: 41) suggest that '[t]here is no need to make a subtle distinction in the taxonomy unless cue phrases exist that reflect it.' The statement assumes that cue phrases are the only indicators of different discourse relations. Signalling, however, is not confined to discourse connectives. Other signals, such as mood and modality in the clause are present. For example, a question (interrogative mood) is a potential signal for a Solutionhood relation (Taboada, 2004b). Non-finite clauses, in some cases, indicate the presence of a Circumstance relation. Example (4), from the RST corpus,¹⁰ shows a Circumstance relationship between spans 1 and 2–5. The satellite, *insisting that they are protected by the Voting Rights Act*, is signalled as such by the non-finite form of the verb, *insisting*.

- (4) [1] Insisting that they are protected by the Voting Rights Act, [2] a group of whites brought a federal suit in 1987 [3] to demand that the city abandon at-large voting for the nine-member City Council [4] and create nine electoral districts, [5] including four safe white districts.

Scott and de Souza (1990) identify some instances of Elaboration through the presence of a relative clause. The head noun denotes the entity being elaborated on, and the relative clause is the elaboration. In other relations, tense and mood are also markers. Scott and de Souza point out that in Brazilian Portuguese, the satellite of some Concession relations (those marked with *embora*) must have subjunctive mood.¹¹

Marcu (2000a) uses connectives to identify some relations. For other relations, such as Elaboration, Joint and Background, he uses lexical chains and cohesive relations (Halliday and Hasan, 1976). Marcu, following previous discussions of the relationship between coreference and cohesion (Harabagiu and Maiorano, 1999), assumes that Elaboration or Background holds between two text spans that discuss the same entities (or entities in subordinate and superordinate relations), whereas Joint holds between segments that contain different entities.

Intonation as a signal of relations is treated in depth by den Ouden (2004). She found that pause duration and pitch were strong indicators of the RST

and about how the impression of coherence is created. These questions deal with foundational issues of the linguistics of communication. The identification of symbolic meaning, by any definition, as the principal basis of communication using

Kittredge, 1993). Sanders and Spooren (1999) refer to this as the 'multi-level thesis'. They argue that a 'strong' multi-level thesis (where every relation is multi-level) is wrong, because there exist clear cases of relations that are either exclusively informational or exclusively presentational. Relations such as Joint, List, or Sequence, are mostly presentational (Goutsos, 1996). We believe that multiple analyses are possible in certain situations, and the formulation of the theory does not preclude them, although they are not encouraged for every relation.

4. Problems with RST analysis

The discussion above associates with RST a descriptive task. Causal accounts may tell why text coherence or RST analyses are found, but these accounts are outside of RST itself. In the search for formal models of human language use, RST is suggestive, but not definitive. There is, however, an issue whether RST analyses should be regarded as worthy of trust, as representative of significant patterns in text that has been analysed. This can be factored for convenience into two types of issues, having to do with judging the analysts and judging the analysis process.

4.1 JUDGING THE ANALYSTS

The judgment issue is restricted to those who analyse text for the purposes of linguistic description. Practising RST analysis, even fairly unskilled analysis, can be helpful in learning writing skills, in guiding attention through a text in a thorough way, in discovering organizational flaws in text, and, of course, in learning RST analysis. Analysis for such purposes is expected to show flaws and divergences that other analysts would not produce. That is normal.

For linguistic description, it was clear even during the early development of RST that practice at analysis led to stability and clarity of judgment. RST analysis of a text does not involve fundamentally different interpretive skills than reading the text, but RST is an explicit, selective, structured way of representing interpretations explicitly. The process of making the interpretations explicit must be learned. Like skill with a musical instrument, use of the RST instrument responds to training. Each analyst also comes with professional and common cultures that supply expectations and influence interpretation.

Judgments as to which relation applies in a particular context are always plausibility judgments. This means that the analyst considers it possible that a particular relation was in the text creator's mind when producing the text. Note that this implies that the relation was *plausibly present*¹³ from the outset, thus implying that the analyst's task is one of redisccovery, not of original discovery of the relation.

There are various ways in which analysts can judge their own skills. Perhaps the best is for analysts to train in groups, with each member analysing sets of shared texts alone, and then comparing and defending (or merging) different analyses.

Several projects have studied analysts' performance in the analysis of discourse, both monologue and dialogue. While there has been little study of analysis using so-called 'Classical RST', the initial set of relations, experience with various near and far relatives of RST analysis is indicative of what is possible. One test that used the RST framework without alteration was a project at the Technical University of Eindhoven, in The Netherlands, preparing for a study on discourse structure and prosody. Six experienced users of RST each analyzed a somewhat complex news report (den Ouden et al., 1998). Their conclusion was that there was high consistency among the analysts, and that disagreements were small in number and could be resolved. Den Ouden (2004) also reports high agreement among judges in RST analyses, higher than in analyses carried out with the Grosz and Sidner (1986) method.

4.2 JUDGING THE ANALYSIS PROCESS

The judgment that RST analysis can be reliable has no inherent significance. There are many processes that can be applied to monologue texts, very reliably, which tell us nothing about those texts. RST analysis is claimed to be significant because of alleged connections to human communication, and in particular to the process of reading. Analysts are reading written texts, a thoroughly subjective process.

Although we may recognize that untrained RST analysts do poorly, they all come to the task as highly trained readers. There is no technical controversy about whether reading can be made reliable in terms of understanding, recall and interpretation. It can, but it cannot be made perfectly reliable, and it takes extensive practice to become a very reliable reader. For RST, as for reading, the task can be decomposed into more elementary subtasks that can be judged separately: dividing text into units; aggregation of units and spans into larger ones; discerning which relation definitions apply to a pair of spans; judging the author's purposes; deciding which span was more central to the author's local purpose. Each of these tasks can be identified, studied, and tested separately. Some studies have performed a number of those tests (Burstein et al., 2001; Carlson et al., 2001a; Marcu et al., 1999a; Teufel et al., 1999). Of course, no array of such tests will establish the significance of RST analysis, but such tests can diagnose the sources of disagreement between analyses, and can lead to improvements to the overall scheme.

One proposal for analysis is the use of underspecified relations. Reitter and Stede (2003) propose an XML format for annotating discourse trees. The innovative aspect of the annotation is that it allows for underspecification of relations. Underspecification was already proposed by Schilder (2002), in an SDRT-based system (Asher and Lascarides, 2003), and it allows for certain aspects of the analysis to be left blank. For instance, nuclearity may be identified easily, but the particular relation between nucleus and satellite can be left underspecified. Similarly, the scope of certain text spans may not be clear until more text is processed. Reitter and Stede (2003) mention the scope problem in relation to sentences starting with adverbials such as *On the other hand, ...*, where the end of the span could be that sentence, or a larger group of sentences.

Underspecification is desirable not only when humans are performing the analysis, but also when we are dealing with automatic parsing, the automatic process of assigning rhetorical structure to a text. In parsing rhetorical relations, further information may be available as analysis progresses, or other modules are called after initial parsing. This annotation method allows for input of spans and relations as they are decided, leaving room for decisions to be made later on in the process. In general, the method may be suitable for cases, automatic or not, when certain decisions can be made fairly quickly, and other decisions – specific relation, scope – can be left unspecified, maybe to be made later.

5. Other relations in text

RST is not the only theory concerned with relations in text. We have already mentioned a number of efforts in similar directions, sometimes diverging, sometimes with intermittent convergences.¹⁴ We would like to explore here a few ways in which relations are interpreted, and points of convergence of other theories with RST.

The comparison most often drawn is between RST and Grosz and Sidner's (1986) account of discourse structure. Grosz and Sidner (henceforth G&S) integrate intentional structure with focus of attention in discourse, relating both to discourse structure. The intentional structure is a reflection of the speaker's goals, meant to achieve some change in the hearer's cognitive status. Intentions can be structured hierarchically, with two possible relations between intentions: dominance and satisfaction-precedence. It has been suggested that RST and G&S are comparable because rhetorical relations are intentional in nature, thus corresponding to G&S's intentional structure. Moser and Moore (1996) propose that nucleus-satellite relations in RST be considered as dominance relations in G&S. This is possible by introducing the notion of a core, the segment that dominates another segment (thus equivalent to the nucleus in RST). One way in which the two theories differ is in the number of relations they postulate. As we have pointed out, a smaller set is not always desirable. When performing G&S analyses, we often find a need for more detailed description than a mere dominance relation. For instance, it is more informative to say that 'segment B is evidence for segment A' than to say 'segment A dominates segment B'. Given intuitive descriptions such as the former, a detailed formal account for each relation, such as RST's, is more precise.

A similar argument on generality applies to relations postulated within Relevance Theory (Sperber and Wilson, 1995). In Relevance Theory, propositions can be in a relation to each other, a relation of relevance. In general, any communicative act carries a presumption of relevance. Relevance is partly derived from context, so when a new utterance is produced after another utterance, the second one will be interpreted as relevant to the first (and to the general context).

Given the hearer's assumption that the speaker has aimed at optimal relevance, it is not surprising that she will assume that an utterance which is part of a text or discourse can be interpreted as somehow following on from the preceding utterances. (Blakemore, 1992: 134)

Discourse markers, in this theory, act as semantic constraints on the types of implicatures a hearer can draw from what the speaker says (Blakemore, 1987, 2002).

In summary, neither Grosz and Sidner nor Relevance Theory seem to say anything further than 'there is some relation'. The specific type of relation is left

through the hierarchical structure of intentions underlying discourse structure in Grosz and Sidner's approach.

There have been attempts at integrating Grosz and Sidner's theory with RST. Moser and Moore (1996) propose that the two theories are fundamentally close, and that differences are minor. The desire to merge both theories stems from an argument that both intentions and rhetorical relations are necessary to generate coherent text (Moore and Paris, 1993). A formalization of the merging is proposed in a coding scheme, called Relational Discourse Analysis (RDA), proposed by Moser and Moore (1995), and used in their case to identify the features that predict discourse marker selection and placement in text generation. RDA has also been used in the coding of tutorial dialogues (Poesio et al., 2002). See also Moore and Pollack (1992).

Other theories of discourse propose a hierarchical structure, with relations among parts. Such is the case of the Geneva pragmatics school (Roulet, 1995; Roulet et al., 1985) or Argumentation Theory (Anscombe and Ducrot, 1983). There is, in both theories, an emphasis on discourse markers, and on how they signal some of the relations.¹⁵

There exist other computationally-oriented efforts to parse structure in discourse. Polanyi (1988, 2001), for instance, has proposed a theory of discourse structure, in which discourse is composed of discourse constituent units (DCUs), related to each other recursively through three different relations: coordination, subordination and binary, sometimes with the presence of an operator (discourse marker). The relations also capture spoken phenomena: repair, for instance, is a type of binary relation. Emphasis is placed on the well-formedness of discourse structures, analogous to constraints on well-formedness of trees that represent the syntactic structure of a sentence.

Veins Theory (Cristea et al., 1998, 2000; Ide and Cristea, 2000) studies the effect of discourse structure on anaphora, based on insights from Centering Theory (Grosz et al., 1995). It is relevant here because it relies on RST for the classification of discourse structure relations (although more emphasis is given to the nucleus-satellite structure than to the type of relation that holds). The theory identifies 'veins' over RST trees, using the nucleus-satellite distinction in RST. Veins define the *discourse structure* for each referring expression.

Discourse Representation Theory (DRT) (Kamp and Reyle, 1993), and its derivative, Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003; Lascarides and Asher, 1993) include a number of rhetorical relations. Some of DRT's and SDRT's rhetorical relations coincide in name with RST (Elaboration), and others are different (Narration). Definitions and applications of the relations are nonetheless quite different. DRT is based on model-theoretical approaches to language (e.g. Montague grammar) and formal semantics (Heim, 1982). Therefore, it is concerned with the conditions under which sentences are true, and the rules used to derive interpretations from the syntactic structure of sentences in context. Asher and Lascarides (2003) provide

an inventory of rhetorical relations, including relations that account for dialogue phenomena (e.g. Question Elaboration, Correction, and Question Answer Pair). One important development is how relations in DRT and especially SDRT have been linked to and used to explain a number of other phenomena: tense, anaphora, VP ellipsis, bridging inferences, implicatures, and presuppositions (Asher and Lascarides, 2003; Lascarides and Asher, 1993).

Other approaches that are very close to RST are within Systemic Functional Linguistics (SFL). Halliday and Hasan (1976) proposed a number of conjunctive relations as part of a general theory of cohesion in text: the types vary slightly: additive, adversative, causal and temporal in Halliday and Hasan (1976); additive, comparative, temporal and consequential in Martin (1992). In both accounts, there is a distinction between *textual*, or experiential relations, and more rhetorical or interpersonal relations, labelled *textual*, which 'obtain in the organization of the text itself rather than the organization of the world the text

enduring contribution of his research. For a description of Bill's research, and a tribute to his life, see the piece written by Christian Matthiessen for *Computational Linguistics* (Matthiessen, 2005).

2. The Web of Science (available through some university libraries), the Google Scholar (scholar.google.com) and the CiteSeer (citeseer.ist.psu.edu) figures are as of 5 July 2005.
3. The name Rhetorical Structure Theory has been criticized in the literature by three groups. One complains that *Rhetorical Structure Theory* is never related to classical rhetoric. A second suggests that a non-standard meaning of *Structure Theory* is being used. The third group points out that RST is thoroughly descriptive and does not put forth a causal *Theory*. All of these are commendable, valid arguments. The name of RST was chosen as a phrasal unit, distinctively best out of a collection of suggestions.
4. Twenty-three relations proper, and one schema, Joint.
5. This view of coherence implicitly ties RST to communication, in contrast to various notions of text structure. Invoking the author's purpose, which RST does not only for the whole text, but for every identified span, makes RST 'structure' quite different from grammatical structure.
6. For example, Knott (1998), which assumed that RST permits only trees with adjacency, and showed that adjacency should not be assumed. There are, however,

- Pragmatics and Semantics: A New Approach to the Theory of Language*, pp. 247–78. London: Taylor and Francis.
- Anscombe, J.-C. and Ducrot, O. (1983) *Les Énoncés à Double Sens*. Brussels: Pierre Mardaga.
- Asher, N. and Lascarides, A. (1994) 'Intentions and Information in Discourse', in J. Pustejovsky (ed.) *Proceedings of the 32nd Meeting of the Association for Computational Linguistics (ACL'94)*, pp. 34–41, Las Cruces, NM.
- Asher, N. and Lascarides, A. (2003) *Language Cognition*. Cambridge: Cambridge University Press.
- Bateman, J. and Delin, J. (2005) 'Rhetorical Structure Theory', in *Encyclopedia of Language and Linguistics*, 2nd edn. Oxford: Elsevier.

- Carlson, L., Marcu, D. and Okurowski, M.E. (2002) *RST Discourse Trees*. Philadelphia, PA: Linguistic Data Consortium.
- Cassell, J., Nakano, Y., Bickmore, T.W., Sidner, C.L. and Rich, C. (2001) 'Non-verbal Cues for Discourse Structure', *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 17–19, Toulouse, France.
- Chafe, W. (1996) 'Beyond Beads on a String and Branches on a Tree', in A.E. Goldberg (ed.) *Coherence and Discourse Structure in Language*, pp. 49–65. Stanford, CA: CSLI.
- Cohen, P.R., Morgan, J.L. and Pollack, M.E. (eds) (1990) *Discourse Processing*. Cambridge, MA: MIT Press.
- Connor, U. (2002) 'New Directions in Contrastive Rhetoric', *Text* 36(4): 493–510.
- Cornish, F. (1989) 'Discourse Structure and Anaphora: Written and Conversational English', *Language* 79(2–3): 229–43.
- Cristea, D., Ide, N. and Romary, L. (1998) 'Veins Theory: A Model of Global Discourse Cohesion and Coherence', *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics*, pp. 17–19, Montreal, Canada.
- Cristea, D., Ide, N., Marcu, D. and Tablan, V. (2000) 'Discourse Structure and Coreference: An Empirical Study', *Proceedings of the 18th International Conference on Computational Linguistics (COLING '00)*, Vol. 1, pp. 208–14, Saarbrücken, Germany.
- Dale, R. (1991) 'The Role of Punctuation in Discourse Structure', *Proceedings of the AAAI Fall Symposium on Discourse Structure and Natural Language Processing*, pp. 13–14, Asilomar, CA.
- Degand, L. and Sanders, T. (2002) 'The Impact of Relational Markers on Expository Text Comprehension in L1 and L2', *Research in Applied Linguistics* 15(7–8): 739–58.
- Degand, L., Lefèvre, N. and Bestgen, Y. (1999) 'The Impact of Connectives and Anaphoric Expressions on Expository Discourse Comprehension', *Discourse Processes* 1(1): 39–51.
- den Ouden, H. (2004) 'Prosodic Realizations of Text Structure', unpublished PhD dissertation, University of Tilburg.
- den Ouden, H., Noordman, L. and Terken, J. (2002) 'The Prosodic Realization of Organisational Features of Text', *Proceedings of the 2002 International Conference on Computational Linguistics*, Aix-en-Provence, France.
- den Ouden, H., van Wijk, C., Terken, J. and Noordman, L. (1998) *Research on User-System Interaction*, Technical University of Eindhoven.
- Eggins, S. and Martin, J.R. (1997) 'Genres and Registers of Discourse', in T.A. van Dijk (ed.) *Discourse and Society: Proceedings of the 1996 International Conference on Discourse and Society*, pp. 230–56. London: Sage.
- Fox, B.A. (1987) *Discourse Structure: A Theory of Text Organization*. Cambridge: Cambridge University Press.
- Fraser, B. (1999) 'What are Discourse Markers?', *Journal of Pragmatics* 31: 931–52.
- Fries, P.H. (2004) 'The Textual Metafunction as a Site for a Discussion of the Goal of Linguistics and Techniques of Linguistic Analysis', paper presented at the 31st International Systemic Functional Congress, Kyoto, Japan.
- Gil, Y. and Ratnakar, V. (2002) 'TRELIS: An Interactive Tool for Capturing Information Analysis and Decision Making', in K. Aho, E. Alami, and K. Alami (eds) *Proceedings of the 2002 International Conference on Computational Linguistics*, pp. 37–42.
- Goutsos, D. (1996) 'A Model of Sequential Relations in Expository Text', *Text* 16

Texts', in O. Rambow (ed.)

- Knott, A. (1996) 'A Data-Driven Methodology for Motivating a Set of Coherence Relations', unpublished PhD dissertation, University of Edinburgh.
- Knott, A. (1998) 'Similarity and Contrast Relations and Inductive Rules', *Proceedings of ACL*, *Volume 1: Distant Reading and Distant Reading Methods*, pp. 54–7, Montréal, Canada.
- Knott, A. and Dale, R. (1994) 'Using Linguistic Phenomena to Motivate a Set of Coherence Relations', *Distant Reading Proceedings* 18(1): 35–62.
- Knott, A. and Dale, R. (1996) 'Choosing a Set of Coherence Relations for Text Generation: A Data-driven Approach', in G. Adorni and M. Zock (eds) *Text Generation: A Data-driven Approach: A Data-driven Approach to Text Generation*, pp. 47–67. Berlin: Springer.
- Knott, A. and Mellish, C. (1996) 'A Feature-based Account of the Relations Signalled by Sentence and Clause Connectives', *Linguistic Inquiry* 39(2–3): 143–83.
- Knott, A. and Sanders, T. (1998) 'The Classification of Coherence Relations and their Linguistic Markers: An Exploration of Two Languages', *Journal of Pragmatics* 30: 135–75.
- Knott, A., Oberlander, J., O'Donnell, M. and Mellish, C. (2001) 'Beyond Elaboration: The Interaction of Relations and Focus in Coherent Text', in T. Sanders, J. Schilperoord and W. Spooren (eds) *Text Generation: A Data-driven Approach to Text Generation*, pp. 181–96. Amsterdam and Philadelphia, PA: John Benjamins.
- Korelsky, T. and Kittredge, R. (1993) 'Towards Stratification of RST', *Proceedings of ACL*, *Volume 1: Distant Reading and Distant Reading Methods*, pp. 52–5, Ohio State University.
- Kruijff-Korbayová, I. and Webber, B. (2000) 'Presuppositional Interpretation of Causal and Additive "Although"', *Proceedings of the Meeting of the Association for Computational Linguistics*, Groningen: University of Groningen.

- Generation to Study Discourse Cue Usage', *Proceedings of AAAI Symposium on the Applications of Artificial Intelligence to Discourse Generation*, pp. 92–8, Stanford, CA.
- Moser, M. and Moore, J.D. (1996) 'Towards a Synthesis of Two Accounts of Discourse Structure', *Computational Linguistics* 22(3): 410–19.
- Nicholas, N. (1994) 'Problems in the Application of Rhetorical Structure Theory to Text Generation', unpublished Master's thesis, University of Melbourne.
- Noordman, L., Dassen, I., Swerts, M. and Terken, J. (1999) 'Prosodic Markers of Text Structure', in K. van Hoek, A. A. Kibrik and L. Noordman (eds) *Discourse Structure: A Cognitive and Computational Approach*, pp. 131–48. Amsterdam and Philadelphia, PA: John Benjamins.
- O'Brien, T. (1995) 'Rhetorical Structure Analysis and the Case of the Inaccurate, Incoherent Source-hopper', *Artificial Intelligence* 16(4): 442–82.
- Poesio, M. and Di Eugenio, B. (2001) 'Discourse Structure and Anaphoric Accessibility', in I. Kruijff-Korbayova and M. Steedman (eds) *Proceedings of ESSLI 2001*, pp. 1–11. Helsinki, Finland.
- Poesio, M., Di Eugenio, B. and Keohane, G. (2002) *Discourse Structure Analysis: A Technical Note* (NLE Technical Note No. TN-02-02). Colchester: University of Essex.
- Poesio, M., Stevenson, R., Di Eugenio, B. and Hitzeman, J. (2004) 'Centering: A Parametric Theory and its Instantiations', *Computational Linguistics* 30(3): 309–63.
- Polanyi, L. (1988) 'A Formal Model of the Structure of Discourse', *Journal of Pragmatics* 12: 601–38.
- Polanyi, L. (2001) 'The Linguistic Structure of Discourse', in D. Schiffrin, D. Tannen and H.E. Hamilton (eds) *The Handbook of Discourse Analysis*, pp. 265–81. Malden, MA: Blackwell.
- Ramachandran, V.S. and Anstis, S.M. (1986) 'The Perception of Apparent Motion', *Scientific American*, October: 102–08.
- Rambow, O. (ed.) (1993) *Proceedings of ACL*, pp. 1–11. Columbus, OH: ACL.
- Redeker, G. (1990) 'Ideational and Pragmatic Markers of Discourse Structure', *Journal of Pragmatics* 14: 367–81.
- Redeker, G. (1991) 'Review Article: Linguistic Markers of Linguistic Structure', *Linguistics* 29(6): 1139–72.
- Redeker, G. (2000) 'Coherence and Structure in Text and Discourse', in W. Black and H. Bunt (eds) *Advances in Discourse Processing: A Cognitive and Computational Approach*, pp. 233–63. Amsterdam and Philadelphia, PA: John Benjamins.
- Reed, C. and Long, D. (1997) *Generating a Paragraph from a Discourse Structure* (Technical Report No. RN/97/157). London: Department of Computer Science, University College London.
- Reitter, D. and Stede, M. (2003) 'Step by Step: Underspecified Markup in Incremental Rhetorical Analysis', *Proceedings of EACL 2003*, pp. 1–11. Budapest, Hungary.
- Rösner, D. (1993) 'Intentions, Rhetoric, or Discourse Relations? A Case from Multilingual Document Generation', in O. Rambow (ed.) *Proceedings of ACL*, pp. 106–09. Ohio State University: ACL.
- Rösner, D. and Stede, M. (1992) 'Customizing RST for the Automatic Production of Technical Manuals', in R. Dale, E. Hovy, D. Rösner and O. Stock (eds) *Advances in Artificial Intelligence*, pp. 119–214. Berlin: Springer.
- Rossari, C. (2001) 'Les relations de discours: Approches rhétoriques, approches pragmatiques et approches sémantiques', *Revue de Linguistique* XXIII(1): 59–72.
- Roulet, E. (1995) 'Geneva School', in J. Verschueren, J.-O. Östman and J. Blommaert (eds) *Pragmatics*, pp. 319–23. Amsterdam and Philadelphia, PA: John Benjamins.

- Roulet, E., Auchlin, A., Moeschler, J., Rubattel, C. and Schelling, M. (1985) *L'écrit*. Bern: Peter Lang.
- Sanders, T. (1986) 'De invloed van globale tekstinesschappen op het begrijpen en onthouden van teksten' ('The Influence of Global Properties of Text on Understanding and Remembering Text'), unpublished master's thesis, Tilburg University.
- Sanders, T. and Noordman, L. (2000) 'The Role of Coherence Relations and their Linguistic Markers in Text Processing', *Discourse Processes* 29(1): 37-60.
- Sanders, T. and Spooren, W. (1999) 'Communicative Intentions and Coherence Relations', in W. Bublitz, U. Lenk and E. Ventola (eds) *Communicative Structures in Discourse: H.A. Coenen - H.A. Dörmann*.

Preferred papers for *Discourse Studies*

Contributions to *Discourse Studies* should satisfy the following criteria:

1. Systematic discourse analysis. *Discourse Studies* (DS) is primarily a discourse analytical journal. That is, articles should provide a detailed, systematic and theoretically based analysis of text or talk. It is insufficient to merely quote, summarize or paraphrase such discourse. Articles are preferred that focus on specific structures or strategies of discourse that are not self-evident to the casual reader. These may include grammatical, stylistic, rhetorical, narrative or argumentative structures; cognitive processes and mental representations; pragmatic, conversational or interactional dimensions of socially situated talk; properties of non-verbal activity (e.g. gestures), images or other graphic elements, among many other properties of communicative events.

Discourse Studies does not publish exclusively theoretical papers, but each paper should feature a prominent theoretical section and a critical review of the relevant literature as a foundation for empirical research. Theoretical notes or short discussion pieces are welcome for the DS Forum section. It goes without saying that both theory and analysis should make an original contribution to the field.

Articles that specifically address the relations between discourse structures, on the one hand, and social and political structures, phenomena and relevant issues, on the other hand, are particularly welcome for DS's companion journal *Discourse & Society*.

2. A sizeable corpus of data. Articles are preferred that are based on a sizeable corpus of interesting texts or talk collected by the author(s) themselves, and not merely on a single or few discourses. Authors are expected to have a thorough knowledge of, and experience with, the corpus, domain or genre of discourse being analysed, for instance as a result of an extended research project, so as to facilitate empirical generalizations and broader insights than those based in one or just a few examples. Analyses should be illustrated by several extracts quoted in the text.

3. Multidisciplinary, multicultural, international. The study of discourse takes place in several disciplines, in many countries and by women and men from many different cultural backgrounds. *Discourse Studies* highly values this diversity and particularly invites contributions which reflect such diversity in their authorship, theories, methods, data and the use of scholarly literature.

4. Accessibility. *Discourse Studies* aims to be accessible to readers from a broad range of disciplines, and of various levels of specialization and expertise, especially including students. For theoretical, methodological, pedagogical and social reasons, therefore, contributions should be well-organized, have a clear style, avoid esoteric jargon, and explain unfamiliar or new technical concepts.