

Rhetorical Structure Theory: Looking Back and Moving Ahead^{*}

Maite Taboada and William C. Mann¹

mtaboada@sfu.ca

January 24, 2006

Abstract

Rhetorical Structure Theory has enjoyed continuous attention since its origins in the 1980s. It has been applied, compared to other approaches, and also criticized in a number of areas in discourse analysis, theoretical linguistics, psycholinguistics, and computational linguistics. In this paper, we review some of the discussions about the theory itself, especially addressing issues of the reliability of analyses and psychological validity, together with a discussion of the nature of text relations. We also propose areas for further research. A follow-up paper (Taboada and Mann, to appear) will discuss applications of the theory in various fields.

1 Reviewing Rhetorical Structure Theory

Rhetorical Structure Theory is a theory of text organization created in the 1980s as a result of exhaustive analyses of texts, described in Mann and Thompson (1988). Since then, the theory has enjoyed a large following, especially in computational linguistics, where it is often used to plan coherent text and to parse the structure of texts.

RST was popular since its very beginning. Citations of the 1988 article in journals begin in 1989 (Cornish, 1989). Other technical reports (Mann and Thompson, 1983) are cited even earlier (e.g., Grosz and Sidner, 1986; Hirschberg and Litman, 1987). Web of Science shows 137 citations of the 1988 paper, and Google Scholar displays 525. CiteSeer contains 167 citations to the 1988 paper, and 172 citations of a paper dated 1987². The latter was a technical report (Mann and Thompson, 1987), which was to

^{*} This paper has been accepted for publication in *Discourse Studies* and the final (edited, revised and typeset) version of this paper will be published in *Discourse Studies* 8 (3), 2006 by Sage Publications Ltd. All rights reserved. © Sage Publications Ltd, 2006.

¹ Bill Mann passed away on August 13, 2004, as this paper was close to a final draft. He had been the driving force behind it, and we had worked together on earlier drafts, but I must now assume all responsibility for omissions and inaccuracies in this version. I only hope that the paper will serve to highlight the enduring contribution of Bill's 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). [MT]

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 July 5, 2005.

appear in a volume edited by Livia Polanyi. The book was never published, but the report was available from the Information Sciences Institute.

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 issues that have been repeatedly mentioned as particularly interesting or troublesome. As for applications, we summarized a number of areas, and significant projects in those. In this paper, we concentrate on theoretical aspects of RST, and we leave a review of its applications for a different paper (Taboada and Mann, to appear). We will have, no doubt, missed important discussions and applications of the theory. Our efforts aim at being representative rather than exhaustive.

The long and fruitful history of RST may make it appear static and fixed. We believe it has a bright future, with many possibilities for expansion and modification. As a consequence, we tried not only to review the past, but also to point out some avenues for further research. Much is still to be resolved and explored. We hope that this paper spurs discussion and investigation of some of those aspects.

We start the paper with a concise introduction to RST in Section 2. Section 3 addresses some of the central issues in RST, examining each set of problems and criticisms and pointing to solutions already proposed and to avenues for further exploration. Section 4 discusses the analysis process, and the reliability and validity of RST analyses. Section 5 is a brief overview of other approaches to coherence and relations in text, and their relationship to RST. Finally, Section 6 provides conclusions and points to areas that could be developed in the future. The applications paper (Taboada and Mann, to appear) will contain the most significant applications of the theory and a list of resources related to RST.

2 RST: A brief introduction

Rhetorical Structure Theory (RST) is about how text works. It is a descriptive linguistic approach to a range of phenomena in the organization of discourse. RST³ was developed

³ The name Rhetorical Structure Theory has been criticized in the literature by three groups. One complains that *Rhetorical* is never related to classical rhetoric. A second suggests that a non-standard meaning of

in the 1980s at the Information Sciences Institute of the University of Southern California by a group of researchers interested in Natural Language Generation: William Mann, Christian Matthiessen and Sandra Thompson, with input from Cecilia Ford, Barbara Fox and Peter Fries. It was 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, to appear).

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 web site (Mann, 2005), or the published description of RST (Mann and Thompson, 1988). The web site 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 *coherence relations*, *discourse relations* or *conjunctive relations* in the literature. Asher and Lascarides (2003) use the term *rhetorical 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,

Structure 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.

3. the visual system resolves confusion
4. by applying some tricks that reflect a built-in knowledge of properties of the physical world.

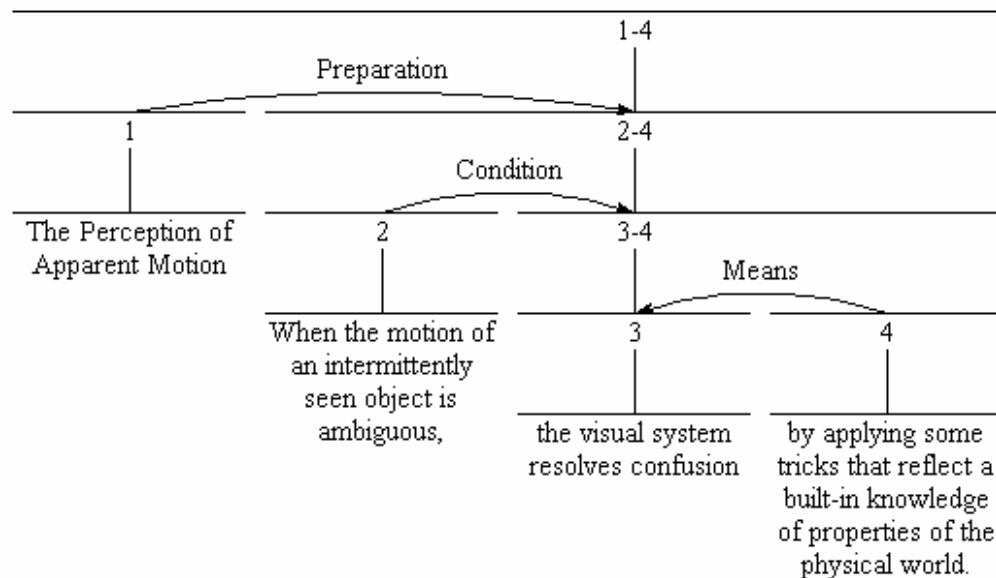


Figure 1. Diagram of an RST analysis

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 paper includes 24 relations⁴, 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 secondary. For instance, in an Elaboration relation, the nucleus is considered to 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 web site. 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.

⁴ Twenty-three relations proper, and one schema, Joint.

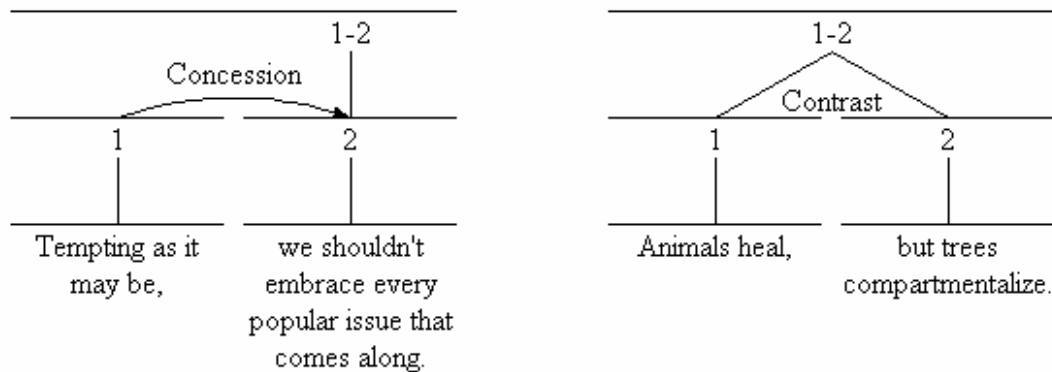


Figure 2. Concession and Contrast relations

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 of the article).

2.4 Why RST?

A question one could ask at this point is why we would want to have a theory such as RST. There are several answers, in the linguistic and the computational side of the theory. First of all, from a linguistic point of view, RST proposes a different view of text organization than most linguistic theories, and a more complete one than most theories of discourse. Second, RST points to a tight relation between relations and coherence in text, thus constituting a way of explaining coherence. One view of text coherence places it in contrast with the presence of non-sequiturs, elements of a text that do not seem to belong, i.e., elements that have no evident purpose. If an RST diagram is a connected whole, with every unit of the text linked into the diagram somehow, then the analysis demonstrates

how the text can be seen as coherent. The possibility of finding an RST diagram for texts is strongly correlated with subjective judgments that texts are coherent⁵. Third, from a computational point of view, it provides a characterization of text relations that has been implemented in different systems, and for applications as diverse as text generation and summarization.

3 Central issues in RST

There has been a surprising diversity of views in the literature concerning how RST characterizes its elements and their use. RST was intended and described as an open system with only a few fixed parts. The lists of relations and schemas were deliberately left open to avoid making claims about any fixed set. Adjacency, tree-shaped analyses, unique analyses without ambiguity, and the possibility of affirming multiple alternatives—all of these have been misunderstood in print. Another cluster of misunderstandings involves the use of independent clauses as the units of analysis. It was, in early descriptions of RST, a suggestion about how circularity could be avoided. This particular suggested selection of units has been taken as a fixed feature of RST, and even as a finding about coherence. And RST has more than once been taken as a draft of an explanatory theory of discourse structure rather than a descriptive system.

Reviewing various misunderstandings item by item, we find that most of them represent digressions that may have done more good than harm. A few, identified in sections below, represent opportunities for improvement. This section examines some of the central notions behind RST and the constituting elements of analysis and analyses.

3.1 What is the goal of RST analyses?

RST's initial goal was the development of a theory that could aid in automatic 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 adequacy and cognitive plausibility (Sanders et al., 1992).

A theory that is descriptively adequate is one that helps characterize the internal structure of texts, producing plausible text structures. Years of text analysis using RST have shown that RST is, indeed, useful to capture the underlying structure of texts. Furthermore, RST has proven to be adequate in computational implementations, in the automatic analysis of texts and in the generation of coherent text. For a summary of these

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.

and other applications in text analysis and in computational linguistics, see Taboada and Mann (to appear).

Cognitive plausibility has been more elusive. It was postulated early on that language users employ RST relations in reading and writing (Mann and Thompson, 1983, 1986), and the topic has been debated at length.

The rest of this section presents a number of issues that contribute to explain those two characteristics: descriptive adequacy and cognitive plausibility. In Section 3.2, we address the process of analysis from the very first step: the segmenting of discourse into units of analysis. Section 3.3 elaborates on cognitive plausibility and the nature of relations, whereas 3.4 takes up the notion of hierarchy and tree structures. One practical issue in the use of RST is the number of relations that can be used, and their possible classifications. This is examined in 3.5. Section 3.6 deals with the problem of unsignalled relations, and how relations can be postulated even when no external signal that they are present is found. Finally, in Section 3.7 we consider the possibility that multiple relations are found in an RST analysis.

3.2 What counts as a unit of analysis?

In order to do analysis, in RST or any comparable text analysis framework, there must be a way to regard the text as consisting of related parts. For RST, the text to be analyzed is divided into units, which then become the minimal elements of the analysis. This division is carried out in a simple way, one intended to be as neutral as possible in influencing the analysis process. Identification of units is done before analysis in order to avoid circularities—analysis depending on the units and unit choices depending on the analysis.

The usual division rule is that each independent clause, along with all of its dependencies of any sort, constitutes a unit. This works reasonably well for many purposes, but it can be problematic for some research goals:

- It misses the fine detail. For long texts, skipping details may be necessary.
- It is tied to the language of the text, and its clause-forming processes. Thus it becomes a source of problems if the text is multilingual, or if it is in a language that is not as clause-centered as English (i.e. where clausal units are not easily established).
- It poses a problem for spoken language, where units are usually considered to be intonation units, and not necessarily independent clauses.

Some researchers have criticized the unit division methods as inadequate, or have pointed out difficulties in establishing units (Marcu, 1998a; Nicholas, 1994; Schauer, 2000; Verhagen, 2001; Wiebe, 1993). We do not believe that one unit division method will be right for everyone; we encourage innovation. For instance, a distinctive innovation affecting unit choices is the work of Rösner and colleagues (Rösner, 1993; Rösner and Stede, 1992). As an alternative to machine translation, they developed multilingual text generation, using RST as a kind of discourse interlingua. Because clauses in one language were not in general clauses in the others (English, French and German in the reported work), the size of the unit was adjusted downward. Another

example of readjusting units is Marcu et al. (2000), an example of Japanese-English machine translation. RST trees are built for the input language, which are then modified to produce adequate trees for the output language. The units have to be adjusted in some cases, to render a more natural output.

RST was initially expected to work well with large unit sizes, such as orthographic paragraphs or subsections, and there are some attempts to make it work in that direction (Granville, 1993; Reed and Long, 1997). Often it does not work well (Marcu et al., 2000). Difficulties arise because in larger units, it is much more likely that there are multiple units that are in some RST relation to items outside of the unit. This sets a practical upper bound on detailed RST analysis. In principle there may be cases, e.g., chapters in a book, in which large units can be found in RST relation. In general, analysis of larger units tends to be arbitrary and uninformative. However, at those levels, other structures are at play: macrostructures (van Dijk, 1980) or constituents of a genre (Eggins and Martin, 1997; Martin, 1984) are probably more informative. Although genre elements, or holistic elements, are mentioned as being outside the scope of RST in Mann and Thompson (1988), no systematic attempt has been made to marry the two approaches. Taboada (2004a) presents an analysis of RST and how certain relations correlate with stages in an analysis of a dialogue genre. A more exhaustive study of different genres would throw light on the relationship between macrostructures or genres and RST structures.

At the opposite end of granularity are analyses of intraclausal relations. In some cases, it may be desirable to analyze relations within the clause, such as Example (1), from Meteor (1993). Here, two actions (flicking the switch; turning the light on) are represented in a single clause, with the verb “caused” indicating a causal relation. We have, however, only one clause, in RST terms, since the complement “(the light) to turn on” would not be a separate unit, and the nominalized verb “flicking” is part of the subject.

(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 relations. Whereas coherence relations realize intentions, there is no one-to-one correspondence between the two.

Den Ouden (2004) performed a very interesting study in which she asked judges to carry out two types of analysis: an intention-based analysis, following Grosz and Sidner (1986), and an RST analysis. Analysts showed better agreement with RST than with the Grosz and Sidner analyses. Den Ouden believes that this is due to the more explicit definitions of RST, and to the more detailed type of analysis required by RST: RST analyses of four texts took judges about seven times longer to complete than a Grosz and Sidner style analysis. It may be the case that intentions do not provide enough detail to explain text structure.

In summary, what counts as a relation must be a clearly established connection between two text spans. Relations, and the hierarchical structure that is captured by

rhetorical relations, help produce the impression of coherence in discourse. There is usually an intention behind expressing the relation, but intentions do not by themselves define relations. Although the presence of a relation is often indicated by lexical or non-lexical cues, relations seem to be present even without clear signalling.

Even though relations are not synonymous with intentions, we still need to question whether they are actually in the minds of language users, or whether they are the product of text analysis. Some accounts (Hobbs, 1979; Knott and Sanders, 1998; Sanders et al., 1993) do consider coherence/rhetorical relations as cognitive entities. In this view, coherence relations are cognitive mechanisms that writers draw upon to join pieces of text together, and that readers recognize when interpreting those pieces.

The issue of psychological validity has, in our view, been too strongly attached to that of type and number of relations. Since we cannot agree on how many relations there are, the argument goes, then it might be the case that those relations are not part of the process of text production or comprehension. This conclusion is obviously not justified. Many fields of study, and many areas of Linguistics, postulate relations without requiring a fixed inventory. Linguists do not attempt to exhaustively enumerate interclausal conjunctions, for example. We take up the issue of relation taxonomy in Section 3.5. In the rest of this section, however, we would like to discuss a number of studies that have shown the importance of relations in the processing of discourse.

Sanders, Spooren and Noordman (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 relations lead to faster processing of discourse segments on-line (Haberlandt, 1982; Sanders, 1986).

Sanders and Noordman (2000) investigated the influence of relation type and explicit marking on reading tasks, testing verification and recall. The hypothesis was that different relations lead to different representations of the text. Their results show that readers had shorter processing times, and better verification and recall when the relation was more complex (a problem-solution relation) than when it was a simple additive relation. This is presumably because more complex relations lead to richer representations. In addition, explicit marking of the relation resulted in faster processing (although it did not affect recall). The conclusion is that text processing relies on relations, resulting in different representations when the relations are different, and that the effect is carried over time to recall. Degand et al. (1999) and Degand and Sanders (2002) further explore the role of connectives in comprehension, showing that they have a positive impact on comprehension tasks after a text has been read.

Another source of evidence that something like RST relations are in the minds of speakers comes from studies in intonation. Den Ouden and colleagues (den Ouden, 2004; den Ouden et al., 2002) observed that certain organizational features of texts corresponded to characteristic prosodic features. They found pause and pitch range correlations with level of embedding of relations, with nuclearity, and with particular

relations: causal relations were associated with shorter pauses than non-causal relations. Noordman et al. (1999) present similar results.

Various reports and papers on RST describe assertions, called *Relational Propositions*, which are produced by the presence of relational structure. The topic has not been prominent, but there is a fairly uncontroversial understanding that such assertions occur. For example, in a text with span C, which presents a claim, connected by an Evidence relation to span E, which presents evidence to support the claim, the relational proposition would assert that the author intends that the evidence will increase the belief of the reader in the claim.

Relational Propositions appear to always be defeasible (deniable) assertions on the part of the author, and their presence and force do not depend on whether the relation is signalled. Precise identification of such propositions would depend on having a formal framework in which to represent propositions corresponding to the text. There is a general understanding that if the text somehow 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 *relational coherence*, as opposed to *entity-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 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 RST⁶. 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 thoughts, language, and actions of others. (Chafe, 1996: 55-56)

More recently, Wolf and Gibson (2004a; 2005) have argued that more powerful data structures than trees are necessary to represent discourse structure. They annotated newspaper texts, using coherence relations based on those proposed by Hobbs (1985) and Kehler (2002). Wolf and Gibson found a large number of crossed dependencies and nodes with more than one parent, which would make it impossible to represent discourse using trees. They propose to use less constrained graphs, although they do not provide full details of the graph structures for large segments of discourse⁸. This could well be an

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, published RST analyses that suspend either adjacency or left-to-right order (Carlson et al., 2001a; Knott, 1998; Knott et al., 2001; Sparck-Jones, 1995; Taboada, 2004b). See also Linde (1979) and Sibun (1992) for differing accounts of tree structures in discourse.

⁷ Created by Mick O'Donnell (www.wagsoft.com).

⁸ A recently released annotated corpus (Wolf et al., 2005) will likely provide more detail.

avenue to explore further. It would be interesting to compare the graph and tree structures, and their interaction with other discourse phenomena, such as discourse markers or anaphoric relations. Graphs are also in line with Chafe's proposals to enhance trees: a flow model in which linear development through time is combined with clustering of ideas into chunks and episodes.

It is simply the case that trees are convenient, easy to represent, and easy to understand. There is, on the other hand, no theoretical reason to assume that trees are the only possible representation of discourse structure and of coherence relations.

3.5 Number and types of relations

One frequent concern is the attempt to define how many rhetorical relations there are—whether for a particular language, or for all languages. This often paired with the attempt to classify relations in certain groups. Mann and Thompson already warned that “no single taxonomy seems suitable” (1988: 256).

The proposal in the 1988 paper was for 24 relations, which can be classified into subject matter (e.g., Elaboration, Circumstance, Solutionhood, Cause, 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).
- Semantic/pragmatic (Sanders et al., 1992, 1993; Schiffrin, 1987; van Dijk, 1977, 1979).
- Ideational/interpersonal. This is a distinction proposed in Systemic Functional Linguistics (Halliday, 1985), and applied to rhetorical relations by others (Maier and Hovy, 1993; Redeker, 2000). The SFL distinction includes a third category, textual, which does not fit clearly against the other dual classifications.
- Experiential/rhetorical (Benwell, 1999).
- Causal/diagnostic, restricted to relations linked by connectives such as *because*, e.g., *The streets are wet because it is raining* versus *It is raining, because the streets are wet* (Traxler et al., 1997).

As it happens with most attempts at classification, the different classifications above represent distinctions that are not fully equivalent. The semantic/pragmatic distinction proposed by van Dijk (1979) separates relations between propositions (semantic) from those between speech acts (pragmatic). This distinction seems to be orthogonal to the subject matter/presentational divide in RST, since both propositions and speech acts can be in the same types of relations. Events denoted by propositions may be related through

Cause, or Background relations; speech acts can be equally related through either type of relation. RST has concentrated on intended effects, more than on how they are achieved, whether it is through presenting relations containing events, facts, propositions or speech acts. This does not mean that the distinction is not worth exploring. Also germane to this discussion is the difference between two types of (semantic) coherence: extensional and intensional (van Dijk and Kintsch, 1983), achieved via relations between either denoted facts and events or between propositions. We are not aware of research that has examined the realization of the different meanings in terms of RST relations, but it would be another interesting pursuit.

Apart from the subject matter/presentational split, another possible classification depends on the signalling of the relations (see Section 3.6). Certain relations are often signalled, often by discourse markers, whereas other relations are rarely or never signalled. Different corpus studies (Knott and Dale, 1994; Taboada, 2004a, 2005) 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 *observability*.

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 concepts are: basic operation (causal or additive); source of coherence (semantic or pragmatic); order of segments (basic or non-basic; basic order is antecedent-consequent, in causal relations); polarity (negative relations link the content of one of the spans to the negation of the other, otherwise relations are positive). Relations can be classified into 12 classes. The classes are constructed by combining the four primitive concepts, and labelling the result. The total number of relations presented in Sanders et al. (1992) is 34. To test the validity of the 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: 30-31) 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 of RST's usefulness. There may never be a single all-purpose hierarchy of defined relations, agreed upon by all. But creating hierarchies that support particular technical purposes seems to be an effective research tactic. The work of Bateman and Rondhuis (1997) is a thoughtful rework of the whole topic, refining the intention structure and the corresponding discourse structure by use of a notion of two strata, linked by realization (one expressing the other). It summarizes and coordinates many efforts that apply the metafunctions of Systemic Functional Linguistics to the realm addressed by RST. It also carefully relates systemic work to other work, producing a synthesis. As the authors note, many of the details remain to be worked out. Nevertheless, the principles and issues that they identify are all worth serious attention.

We reiterate here that the taxonomy of relations is not fixed. A reduction of the current 30 to more general classes is possible, as Sanders and colleagues (Sanders et al., 1992, 1993) or Louwerse (2001) have done. An extension is also feasible, but with a caveat: a larger number of relations will increase the difficulty of manual text analysis, since there is a possible limit to how many relations, with their corresponding constraints and effects, can be managed simultaneously by an analyst.

3.6 Interpreting unsignaled relations

There is an unexpected and still unexplained phenomenon that has come into view with the use of RST. It is the recognizable presence of relations that are, seemingly, not signalled in any explicit way. Relations can be signalled by cue phrases (discourse markers or discourse particles), mood, tense and aspect, or structural characteristics, such as adjacency pairs in conversation. It has been argued (Scott and de Souza, 1990) that all rhetorical relations proposed in the 1988 paper 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, Fraser (1999), Redeker (1990; 1991), Schiffrin (1987; 2001), Blakemore (1987; 1992; 2002), or Louwerse and Mitchell (2003). 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, 2005), of Wall Street Journal articles from the RST corpus (Carlson et al., 2002), shows levels of signalling ranging from 4% of all Summary relations (a total of 75) to over 90% 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% 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% 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% 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 other⁹.

The question we are interested in here is the issue of unsignaled relations. One position is that the fact that unsignaled 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 corpus¹⁰, 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 mood¹¹.

⁹ See Kehler (2002) and Wolf et al. (2004) for a view of the relationship between coherence and cohesion. The argument is that pronoun resolution (cohesive relation) is part of establishing coherence in discourse, as we have discussed in Section 3.3.

¹⁰ This example is also discussed in Taboada(2005).

¹¹ The signalling, however, is probably most strongly conveyed by the discourse marker than by the mood of the verb.

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 structure of read-aloud texts. She tested a number of characteristics of text structure, and looked for intonation correlates. The nucleus-satellite distinction is not signalled by pause duration or pitch, but den Ouden found a correlation with articulation rate: nuclei were read more slowly than satellites. She also found differences between segments that were related through causal and non-causal (additive) relations¹², and longer pauses preceding segments related through presentational relations, as compared to segments in subject matter relations.

Cassell and colleagues (2001) found that some aspects of discourse structure, such as topic changes, were signalled by changes in posture. New discourse segments at the beginning of a turn are likely to be accompanied by posture shifts. When speakers produce the end of a discourse segment at the end of a turn, their posture shifts last longer than when the two ends (of discourse segment and turn) do not co-occur. The authors point out that these findings relate to unit boundaries, but that further research may indicate a relationship between posture and information content of units (and possibly relationships among units, we would add).

Bateman and colleagues have explored the signalling of relations through punctuation and layout for the purpose of natural language generation (Bateman et al., 2001). Preliminary work reported by Dale (1991) addresses the meaning of punctuation and paragraph breaks.

It is clear then that there are a large number of mechanisms to indicate the presence of a relation. The question is not, however, whether relations *can* be signalled, but whether they routinely are. Since the answer is that they can be unsignalled, we need to question how unsignalled relations can be interpreted. And how relations that are signalled can be interpreted *unambiguously*. The most common views of language and communication would suggest that lack of signalling is not possible as a recurrent, widely distributed kind of event. Semiotics and common linguistic models suggest that the symbols (or signs) of texts are the sole source of the particular interpretations that the text can produce. Recognition of RST relations would seem to depend on the reliable presence of signals. Absence of signals can be expected to produce incoherence.

¹² Differences in two different studies were in opposite directions: in a corpus-based study, pauses between causally related segments were shorter; in a controlled study with constructed texts, pauses were found to be longer between causally related sentences, maybe because relations were not marked lexically.

This turns out to be an incorrect expectation. RST analyses of texts from a very diverse range of sources show that the absence of such signals is extremely frequent. The frequency may vary for different text types, but an estimate of over 50% of the relations being unsignaled is representative. The analyses on the RST web site (Mann, 2005), a very diverse collection comprising 187 units, have about 72% of the relations unsignaled. Although unsignaled relational discourse structure is abundant, its significance is not in its frequency. Rather, it is qualitative: the issue of how unsignaled discourse structure is possible at any frequency. Coherence of text and absence of relation signalling need to be considered together.

Several questions arise, about how reliable analysis is possible, how readers and analysts discover relations if the supposedly necessary signals are missing, 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 language, is called into question. One proposal that we have put forth elsewhere (Taboada, 2005) is that expectations about how texts (and possibly conversations) proceed provide enough information to interpret higher-level relations. This brings us back to the relationship between rhetorical relations and the structure of genres in discourse (see Section 3.2). A reader may recognize the last few sentences of a text as a summary of the whole text because he or she is familiar with such texts. This is only a partial answer, and speculative. The issue needs to be addressed systematically, ideally through studies of processing.

3.7 Multiple analyses

RST originally restricted the relation between two text spans (or units) to one single RST relation. If multiple candidates are possible (see Section 4.2, on the analysis process), the analyst is to choose the most plausible relation. Sometimes, however, there seem to be pairs of relations holding in the same place, as in Example (5), a constructed example, where both Elaboration and Antithesis could both be applied. Such combinations can be stated, without creating a new relation, if the analyst simply affirms that two different analyses hold, which is generally an accurate if awkward approach.

(5) This is good juice. It is made from concentrate.

Moore and Pollack (1992) and Redeker (2000), among others, have noted that often there seem to be two relations holding between two text units. Moore and Pollack (1992) establish informational and intentional levels, pointing out that such organization is accounted for in the original publications on RST, in the split between presentational and subject matter relations (see Section 3.5). They use an example, reproduced in (6), with two possible analyses, an Evidence relation, with (6a) as evidence for (6b), or a Cause relation. They provide descriptions of the contexts in which each of those interpretations could occur.

- (6) a. George Bush supports big business.
b. He's sure to veto House Bill 1711.

Most RST analyses, however, are based on the contexts in which texts are produced, and on knowledge of hearer and reader. Therefore, if Example (6) had a context, the

ambiguity would likely not arise. Asher and Lascarides (1994) provide three different contexts that would constrain the interpretation of the example in (6), and Sanders and Spooren (1999) also show that the problem with the example is underspecification or vagueness. They provide five different contexts that make the relation explicit.

Similar claims that two relations may hold between pairs of sentences have been made elsewhere (Andriessen et al., 1996; Asher and Lascarides, 2003; Bateman and Rondhuis, 1997; Martin, 1992: 258-260; Redeker, 1990, 2000), and there are proposals to formalize the use of two levels in analysis (Korelsky and 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 Practical issues in using the theory

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 analyzed. 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 analyze text for the purposes of linguistic description. Practicing 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 *in the writer’s mind*¹³ from the outset, thus implying that the analyst’s task is one of rediscovery, 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 analyzing 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.

Extensive testing of analysts’ agreement is reported in the work of Marcu and colleagues (Marcu et al., 1999a, 1999b). An extensive modified version of RST, including new instructions and 78 relations, was used to test agreement on three different levels: assignment of text spans, assignment of nuclei, and assignment of relations. The report concludes that “... even simple, intuitive definitions of rhetorical relations... and discourse structure can lead to reliable annotation schemata.” (Marcu et al., 1999a: 55). Another finding of this study was that none of the simple sequence schemes for organizing analysis (left-to-right, bottom-up, top-down) was practical at every point. Rather, opportunistic analysis is necessary.

Sanders and van Wijk (1996) describe PISA (Procedure for Incremental Structure Analysis), a procedure for annotating texts incrementally that relies on linguistic cues to postulate the presence of relations. The procedure was especially formulated with annotator reliability in mind. Although no specific discussion of agreement tests and results is provided, it is clear that the method is successful because it provides a clear methodology for postulating relations.

Teufel and colleagues (Teufel et al., 1999) tested a methodology for the annotation of argumentation in research articles. When examining their own performance with a six-page annotation manual, the researchers found their results to be highly consistent and reproducible, with high agreement when they performed the same task six weeks later.

In general, we see this as a problem akin to many other category assignments in text analysis, such as speech act assignment given a list of speech acts (Stolke et al., 2000), or deciding on sentence topic (Byron and Stent, 1998). In all these cases, agreement is

¹³ We discuss RST in conversation in Taboada and Mann (to appear). See also Taboada (2004a, chapter 5).

reached through training and discussion of problematic cases on a test set before actual analysis is performed. Agreement among analysts can be assessed using, among other methods, the kappa coefficient (Carletta, 1996; Siegel and Castellan, 1988), as has been done in many of the studies mentioned above. Fundamental to the success of any analysis process is the creation of solid guidelines, so that the decisions made in the analysis are explicit and reproducible. Most of the corpus annotations efforts have produced manuals to guide the annotation process (Carlson and Marcu, 2001; Wolf et al., 2003).

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 added later.

5 Other approaches to 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 convergences¹⁴. 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).

¹⁴ Bateman and Delin (2005) also provide an interesting discussion of the connections between RST and other levels of linguistic description.

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 to be inferred, in Relevance Theory, or is considered sufficiently described 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 relations¹⁵.

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., 2000; Cristea et al., 1998; 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 *domain of referential accessibility* 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

¹⁵ For a comparison between RST and the Geneva school, see Rossari (2001).

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 *external*, or experiential relations, and more rhetorical or interpersonal relations, labelled *internal*, which “obtain in the organization of the text itself rather than the organization of the world the text describes” (Martin, 1992: 180). Conjunction is based on the presence of conjunctive elements, which are a subset of the general category of discourse markers. In SFL, conjunctive elements are mostly conjunctions and prepositional phrases: *and*, *but*, *so*, *by the way*, *for this reason*, etc. Halliday and Hasan (1976: 321-322) suggest that other classifications of conjunctive relations are possible; e.g., following traditional classifications of subordinate clauses.

6 Conclusions; moving ahead

Reviewing almost twenty years of development and use of RST, the value that we find in RST is in:

- the analyses and consequent understandings of text that are enabled,
- the conceptual structure of relations and coherence, and
- the contribution to a great diversity of work in several fields in which RST is used as a conceptual starting point, far beyond text generation, the initial target.

We have seen how RST provides a bird’s eye view of a text, pinpointing the most important parts, and helping to make explicit some of the implicit relations present. We have also shown how RST has provided a way of defining relations so that they can be applied in different domains and by different researchers. Rhetorical relations provide an explanation for the perceived coherence of most texts.

What is the outlook for RST? It seems to us that it is a healthy, well-tested theory, with a large number of practitioners. It should continue to have currency, maybe in part through innovation and applications in new areas. To guide some of that innovation, and to encourage applications, we would like to point out some open areas of research, partly drawn from the RST web site (Mann, 2005).

A first group of open issues has to do with systematic, large-scale application of RST to texts. Although there are applications in English to corpora of substantial size (Carlson

et al., 2001b), there are not many studies in languages other than English. Annotated corpora, such as the one under construction for German (Stede, 2004) would be useful in exploring realization, frequency, and signalling of rhetorical relations. Comparisons across languages and text types—both synchronically and diachronically—would provide insights into language universals and language change.

RST has been mostly presented as a descriptive tool, and we have discussed some of the problems in endowing the theory with psychological validity. It would be worthwhile to explore how RST can help explain reading and understanding processes in text, and how text is created (especially on-line). A thorough investigation of these issues would touch upon models of language use, memory, and processing.

A third area of research lies in the interface of rhetorical structure and many other discourse phenomena. There have been investigations on the relationship of discourse structure, using RST or cognate theories, and anaphoric relations (Fox, 1987; Poesio and Di Eugenio, 2001); on the interaction of discourse structure, anaphoric relations and thematic roles (Stevenson et al., 2000); and discourse structure and clause-internal phenomena such as VP ellipsis, gapping, or extraction (Kehler, 2002), but further research in those areas is possible. It is also worth exploring how RST interacts with cohesive devices in the language. Fries (2004) has proposed to relate RST to information structure, in the form of Theme-Rheme relations. The distribution and realization of rhetorical relations across different genres, text types and modes deserves attention. We may find that the same relation exhibits different realizations in different genres, or that presence and frequency of certain relations vary in conversation versus written text. In the same spirit we could explore the interface of RST with different semantic theories, speech act theory, or theories of intentionality.

Acknowledgments

The first author was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada and by project MCYT-FEDER BFF2002-02441/XUGA-PGIDIT03PXIC20403PN, from the Ministry of Science and Technology of Spain and the Xunta de Galicia (PI: MLA Gómez-González). We would like to thank Dennis Storoshenko for his help in tracking down RST-related material and bibliographical references; John Bateman, Giuseppe Carenini, Wallace Chafe, Mick O'Donnell and Manfred Stede for comments on an earlier draft; an anonymous reviewer for many suggestions and references that helped focus the paper; and Teun van Dijk for his editorial guidance.

References

- Afantenos, S. D., I. Doura, E. Kapellou and V. Karkaletsis. (2004). Exploiting cross-document relations for multi-document evolving summarization. In *Methods and Applications of Artificial Intelligence, Proceedings of SETN 2004* (pp. 410-419). Berlin: Springer.
- Andriessen, Jerry, Koenraad de Smedt and Michael Zock. (1996). Discourse Planning: Experimental and modeling approaches. In T. Dijkstra and K. de Smedt (Eds.),

- Computational Psycholinguistics: Symbolic and Network Models of Language Processing* (pp. 247-278). London: Taylor and Francis.
- Anscombre, Jean-Claude and Oswald Ducrot. (1983). *L'argumentation dans la langue*. Brussels: Pierre Mardaga.
- Asher, Nicholas and Alex Lascarides. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.
- Asher, Nicholas and Alex Lascarides. (1994). Intentions and information in discourse. *Proceedings of 32nd Meeting of the Association for Computational Linguistics (ACL'94)* (pp. 34-41). Las Cruces, New Mexico.
- Bateman, John and Judy Delin. (2005). Rhetorical structure theory. In *Encyclopedia of Language and Linguistics* (2nd ed.). Oxford: Elsevier. (to appear).
- Bateman, John, Thomas Kamps, Jörg Klein and Klaus Reichenberger. (2001). Towards constructive text, diagram, and layout generation for information presentation. *Computational Linguistics*, 27 (3), 409-449.
- Bateman, John and Klaas Jan Roonhuis. (1997). Coherence relations: Towards a general specification. *Discourse Processes*, 24, 3-49.
- Benwell, Bethan. (1999). The organisation of knowledge in British university tutorial discourse: Issues, pedagogic discourse strategies and disciplinary identity. *Pragmatics*, 9 (4), 535-565.
- Blakemore, Diane. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Blakemore, Diane. (1992). *Understanding Utterances: An Introduction to Pragmatics*. Oxford: Blackwell.
- Blakemore, Diane. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Bloom, Lois, Margaret Lahey, Lois Hood, Karin Lifter and Kathleen Fiess. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, 7, 235-261.
- Bouayad-Agha, Nadjat. (2000). Using an abstract rhetorical representation to generate a variety of pragmatically congruent texts. *Proceedings of the 38th Meeting of the Association for Computational Linguistics (ACL'00), Student Workshop* (pp. 16-22). Hong Kong.
- Bratman, Michael. (1987). *Intentions, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Burstein, Jill C., Karen Kukich, Slava Andreyev and Daniel Marcu. (2001). Towards automatic classification of discourse elements in essays. *Proceedings of 39th Annual Meeting of the Association for Computational Linguistics (ACL'01)*. Toulouse, France.
- Byron, Donna K. and Amanda Stent. (1998). A preliminary model of Centering in dialog. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics (ACL-98)* (pp. 1475-1477). Montréal, Canada.
- Carletta, Jean. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, 22 (2), 249-254.
- Carlson, Lynn, John Conroy, Daniel Marcu, Dianne O'Leary, Mary Ellen Okunowski, Anthony Taylor and William Wong. (2001a). An empirical study of the relation between abstracts,

- extracts, and the discourse structure of texts. *Proceedings of Document Understanding Conference (DUC-2001)*. New Orleans, Louisiana.
- Carlson, Lynn and Daniel Marcu. (2001). *Discourse Tagging Manual*. Unpublished manuscript, <http://www.isi.edu/~marcu/discourse/tagging-ref-manual.pdf>.
- Carlson, Lynn, Daniel Marcu and Mary Ellen Okurowski. (2001b). Building a discourse tagged corpus in the framework of Rhetorical Structure Theory. *Proceedings of 2nd SIGDIAL Workshop on Discourse and Dialogue, Eurospeech 2001*. Aalborg, Denmark.
- Carlson, Lynn, Daniel Marcu and Mary Ellen Okurowski. (2002). RST Discourse Treebank [Corpus]. Philadelphia, PA: Linguistic Data Consortium.
- Cassell, Justine, Yukiko Nakano, Timothy W. Bickmore, Candace L. Sidner and Charles Rich. (2001). Non-verbal cues for discourse structure. *Proceedings of the 41st Meeting of the Association for Computational Linguistics* (pp. 17-19). Toulouse, France.
- Chafe, Wallace. (1996). Beyond beads on a string and branches on a tree. In A. E. Goldberg (Ed.), *Conceptual Structure, Discourse and Language* (pp. 49-65). Stanford, CA: CSLI.
- Cohen, Philip R., Jerry L. Morgan and Martha E Pollack (Eds.). (1990). *Intentions in Communication*. Cambridge, MA: MIT Press.
- Connor, Ulla. (2002). New directions in contrastive rhetoric. *Tesol Quarterly*, 36 (4), 493-510.
- Cornish, Francis. (1989). Discourse structure and anaphora: Written and conversational English. *Lingua*, 79 (2-3), 229-243.
- Cristea, Dan, Nancy Ide, Daniel Marcu and Valentin Tablan. (2000). Discourse structure and coreference: An empirical study. *The 18th International Conference on Computational Linguistics (COLING'00)* (Vol. 1, pp. 208-214). Saarbrücken, Germany.
- Cristea, Dan, Nancy Ide and Laurent Romary. (1998). Veins Theory: A model of global discourse cohesion and coherence. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational Linguistics (ACL-98/COLING-98)* (pp. 281-285). Montréal, Canada.
- Dale, Robert. (1991). The role of punctuation in discourse structure. *Proceedings of AAAI Fall Symposium on Discourse Structure in Natural Language Understanding and Generation* (pp. 13-14). Asilomar, CA.
- Degand, Liesbeth, Nathalie Lefèvre and Yves Bestgen. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design*, 1 (1), 39-51.
- Degand, Liesbeth and Ted Sanders. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, 15 (7-8), 739-758.
- den Ouden, Hanny. (2004). *Prosodic Realizations of Text Structure*. Unpublished Ph.D. dissertation, University of Tilburg, Tilburg, The Netherlands.
- den Ouden, Hanny, Leo Noordman and Jacques Terken. (2002). The prosodic realization of organisational features of text. *Proceedings of Speech Prosody 2002*. Aix-en-Provence, France.
- den Ouden, Hanny, Carel van Wijk, Jacques Terken and Leo Noordman. (1998). *Reliability of Discourse Structure Annotation* (No. 33, Annual Progress Report). Eindhoven: IPO Center for Research on User-System Interaction, Technical University of Eindhoven.

- Eggins, Suzanne and James R. Martin. (1997). Genres and registers of discourse. In T. A. v. Dijk (Ed.), *Discourse as Structure and Process. Discourse Studies: A Multidisciplinary Introduction* (pp. 230-256). London: Sage.
- Fox, Barbara A. (1987). *Discourse Structure and Anaphora: Written and Conversational English*. Cambridge: Cambridge University Press.
- Fraser, Bruce. (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931-952.
- Fries, Peter 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 V. Ratnakar. (2002). TRELLIS: An interactive tool for capturing information analysis and decision making. In *Knowledge Engineering and Knowledge Management, Proceedings* (pp. 37-42).
- Goutsos, Dionysis. (1996). A model of sequential relations in expository text. *Text*, 16 (4), 501-533.
- Granville, Robert Alan. (1993). An algorithm for high-level organization of multi-paragraph texts. *Proceedings of ACL Workshop on Intentionality and Structure in Discourse Relations* (Vol. 1, pp. 19-22). Columbus, Ohio.
- Grasso, F. (2003). Rhetorical coding of health promotion dialogues. In *Artificial Intelligence in Medicine, Proceedings* (Vol. 2780, pp. 179-188).
- Grosz, Barbara J., Aravind K. Joshi and Scott Weinstein. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21 (2), 203-225.
- Grosz, Barbara J. and Candace L. Sidner. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12 (3), 175-204.
- Grosz, Barbara J. and Candace L. Sidner. (1990). Plans for discourse. In P. R. Cohen, J. L. Morgan and M. E. Pollack (Eds.), *Intentions in Communication* (pp. 417-444). Cambridge, MA: MIT Press.
- Haberlandt, Karl. (1982). Reader expectations in text comprehension. In J.-F. Le Ny and W. Kintsch (Eds.), *Language and Comprehension* (pp. 239-249). Amsterdam: North-Holland.
- Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, Michael A. K. and Ruqaiya Hasan. (1976). *Cohesion in English*. London: Longman.
- Harabagiu, Sanda and Steven Maiorano. (1999). Knowledge-lean coreference resolution and its relation to textual cohesion and coherence. *Proceedings of ACL Workshop on The Relation of Discourse/Dialogue Structure and Reference* (pp. 29-38). College Park, Maryland.
- Heim, Irene. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Unpublished Ph.D. Dissertation, University of Massachusetts Amherst.
- Hirschberg, Julia and Diane J. Litman. (1987). Now let's talk about now: Identifying cue phrases intonationally. *Proceedings of 25th Annual Meeting of the Association for Computational Linguistics (ACL'87)* (pp. 163-171). Stanford, CA.
- Hobbs, Jerry. (1979). Coherence and coreference. *Cognitive Science*, 6, 67-90.

- Hobbs, Jerry. (1985). *On the Coherence and Structure of Discourse* (Research Report 85-37). Stanford, CA: CSLI.
- Hovy, Eduard. (1988). Planning coherent multisentential text. *Proceedings of 26th Annual Meeting of the Association for Computational Linguistics (ACL'88)* (pp. 163-169). Buffalo, New York.
- Hovy, Eduard. (1990). Parsimonious and profligate approaches to the question of discourse structure relations. *Proceedings of 5th International Workshop on Text Generation* (pp. 59-65). Pittsburgh, PA.
- Hovy, Eduard, Julia Lavid, Elisabeth Maier, Vibhu Mittal and Cécile Paris. (1992). Employing knowledge resources in a new text planning architecture. In R. Dale, E. Hovy, D. Rösner and O. Stock (Eds.), *Aspects of Automated Natural Language Generation* (pp. 57-72). Berlin: Springer.
- Hovy, Eduard and Elisabeth Maier. (1995). *Parsimonious or Profligate: How Many and Which Discourse Structure Relations?* Unpublished manuscript.
- Hubert, Lawrence. (1977). Kappa revisited. *Psychological Bulletin*, 84 (2), 289-297.
- Ide, Nancy and Dan Cristea. (2000). A hierarchical account of referential accessibility. *Proceedings of the 38th Meeting of the Association for Computational Linguistics*. Hong Kong.
- Kamp, Hans and Uwe Reyle. (1993). *From Discourse to Logic: Introduction to Model Theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. The Netherlands: Dordrecht.
- Károly, Krisztina. (1998). Written text analysis: A multidisciplinary field of study. *Sprachtheorie und germanistische Linguistik*, 8 (1), 71-108.
- Kehler, Andrew. (2002). *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI.
- Kibble, Rodger and Richard Power. (2004). Optimizing referential coherence in text generation. *Computational Linguistics*, 30 (4), 401-416.
- Knott, Alistair. (1996). *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Unpublished Ph.D. dissertation, University of Edinburgh, Edinburgh, UK.
- Knott, Alistair. (1998). Similarity and contrast relations and inductive rules. *Proceedings of ACL Workshop on Discourse Relations and Discourse Markers* (pp. 54-57). Montréal, Canada.
- Knott, Alistair and Robert Dale. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18 (1), 35-62.
- Knott, Alistair and Robert Dale. (1996). Choosing a set of coherence relations for text generation: A data-driven approach. In G. Adorni and M. Zock (Eds.), *Trends in Natural Language Generation: an Artificial Intelligence Perspective* (pp. 47-67). Berlin: Springer.
- Knott, Alistair and Chris Mellish. (1996). A feature-based account of the relations signalled by sentence and clause connectives. *Language and Speech*, 39 (2-3), 143-183.
- Knott, Alistair, Jon Oberlander, Michael O'Donnell and Chris Mellish. (2001). Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord and W. Spooren (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects* (pp. 181-196). Amsterdam and Philadelphia: John Benjamins.

- Knott, Alistair and Ted Sanders. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, 30, 135-175.
- Korelsky, Tanya and Richard Kittredge. (1993). Towards stratification of RST. *Proceedings of ACL Workshop on Intentionality and Structure in Discourse Relations* (pp. 52-55). Ohio State University.
- Kruijff-Korbayová, Ivana and Bonnie Webber. (2000). Presuppositional Interpretation of causal and additive "although". *Proceedings of Making Sense: From Lexeme to Discourse*. Groningen.
- Lascarides, Alex and Nicholas Asher. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16 (5), 437-493.
- Linde, Charlotte. (1979). Focus of attention and the choice of pronouns in discourse. In T. Givón (Ed.), *Syntax and Semantics. Volume 12: Discourse and Syntax* (pp. 337-354). New York: Academic Press.
- Litman, Diane J. and James F. Allen. (1990). Discourse processing and commonsense plans. In P. R. Cohen, J. L. Morgan and M. E. Pollack (Eds.), *Intentions in Communication* (pp. 365-388). Cambridge, MA: MIT Press.
- Louwerse, Max M. (2001). An analytic and cognitive parametrization of coherence relations. *Cognitive Linguistics*, 12 (3), 291-315.
- Louwerse, Max M. and Heather H. Mitchell. (2003). Toward a taxonomy of a set of discourse markers in dialogue: A theoretical and computational linguistic account. *Discourse Processes*, 35 (3), 243-281.
- Maier, Elisabeth and Eduard Hovy. (1991). A metafunctionally motivated taxonomy for discourse structure relations. *Proceedings of 3rd European Workshop on Language Generation*. Innsbruck, Austria.
- Maier, Elisabeth and Eduard Hovy. (1993). Organising discourse structure relations using metafunctions. In H. Horacek and M. Zock (Eds.), *New Concepts in Natural Language Generation* (pp. 69-86). London: Pinter.
- Mann, William C. (2005). *RST Web Site*, from <http://www.sfu.ca/rst>
- Mann, William C., Christian M.I.M. Matthiessen and Sandra A. Thompson. (1992). Rhetorical Structure Theory and text analysis. In W. C. Mann and S. A. Thompson (Eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text* (pp. 39-78). Amsterdam and Philadelphia: John Benjamins.
- Mann, William C. and Sandra A. Thompson. (1983). *Relational Propositions in Discourse* (Technical Report No. ISI/RR-83-115). Marina del Rey, CA: Information Sciences Institute.
- Mann, William C. and Sandra A. Thompson. (1986). *Rhetorical Structure Theory: Description and Construction of Text Structures* (Technical Report No. ISI/RS-86-174). Marina del Rey, CA: Information Sciences Institute.
- Mann, William C. and Sandra A. Thompson. (1987). *Rhetorical Structure Theory: A Theory of Text Organization* (No. ISI/RS-87-190). Marina del Rey, CA: Information Sciences Institute.
- Mann, William C. and Sandra A. Thompson. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8 (3), 243-281.

- Marcu, Daniel. (1996). Building up rhetorical structure trees. *Proceedings of 13th National Conference on Artificial Intelligence* (Vol. 2, pp. 1069-1074). Portland, Oregon.
- Marcu, Daniel. (1997). *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts*. Unpublished Ph.D. dissertation, University of Toronto, Toronto, Canada.
- Marcu, Daniel. (1998a). A surface-based approach to identifying discourse markers and elementary textual units in unrestricted texts. *Proceedings of COLING-ACL Workshop on Discourse Relations and Discourse Markers* (pp. 1-7). Montréal, Canada.
- Marcu, Daniel. (1998b). To build text summaries of higher quality, nuclearity is not sufficient. *Proceedings of AAAI-98 Spring Symposium on Intelligent Text Summarization* (pp. 1-8). Stanford, California.
- Marcu, Daniel. (2000a). The rhetorical parsing of unrestricted texts: A surface based approach. *Computational Linguistics*, 26 (3), 395-448.
- Marcu, Daniel. (2000b). *The Theory and Practice of Discourse Parsing and Summarization*. Cambridge, Mass: MIT Press.
- Marcu, Daniel, Lynn Carlson and Maki Watanabe. (2000). The automatic translation of discourse structures. *1st Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL'00)* (Vol. 1, pp. 9-17). Seattle, Washington.
- Marcu, Daniel, Magdalena Romera and Estibaliz Amorrortu. (1999a). Experiments in constructing a corpus of discourse trees. *Proceedings of ACL Workshop on Standards and Tools for Discourse Tagging* (pp. 48-57). College Park, Maryland.
- Marcu, Daniel, Magdalena Romera and Estibaliz Amorrortu. (1999b). Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. *Workshop on Levels of Representation in Discourse* (pp. 71-78). Edinburgh, UK.
- Martin, James R. (1984). Language, register and genre. In F. Christie (Ed.), *Children Writing: Reader* (pp. 21-30). Geelong, Victoria: Deakin University Press.
- Martin, James R. (1992). *English Text: System and Structure*. Amsterdam and Philadelphia: John Benjamins.
- Matthiessen, Christian M.I.M. (2005). Remembering Bill Mann. *Computational Linguistics*, 31 (2), 161-172.
- Matthiessen, Christian M.I.M. and Sandra A. Thompson. (1988). The structure of discourse and "subordination". In J. Haiman and S. A. Thompson (Eds.), *Clause Combining in Discourse and Grammar* (pp. 275-329). Amsterdam and Philadelphia: John Benjamins.
- Meteor, Marie. (1993). Assumptions underlying discourse relations: Which ones are really there and where are they? *Proceedings of Workshop on Intentionality and Structure in Discourse Relations, ACL* (pp. 82-85). Ohio State University.
- Meyer, Bonnie J. F., David M. Brandt and George J. Bluth. (1980). Use of top-level structure in text: Key for reading comprehension in ninth-grade students. *Reading Research Quarterly*, 16 (1), 72-103.
- Moore, Johanna D and Cécile Paris. (1993). Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19 (4), 651-694.
- Moore, Johanna D and Martha E Pollack. (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics*, 18 (4), 537-544.

- Moser, Megan and Johanna D Moore. (1995). Using discourse analysis and automatic text generation to study discourse cue usage. *Proceedings of AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation* (pp. 92-98). Stanford, California.
- Moser, Megan and Johanna D Moore. (1996). Towards a synthesis of two accounts of discourse structure. *Computational Linguistics*, 22 (3), 410-419.
- Nicholas, Nick. (1994). *Problems in the Application of Rhetorical Structure Theory to Text Generation*. Unpublished Master's thesis, University of Melbourne, Melbourne.
- Noordman, Leo, Ingrid Dassen, Mar Swerts and Jacques Terken. (1999). Prosodic markers of text structure. In K. van Hoek, A. A. Kibrik and L. Noordman (Eds.), *Discourse Studies in Cognitive Linguistics: Selected Papers from the Fifth International Cognitive Linguistics Conference* (pp. 131-148). Amsterdam and Philadelphia: John Benjamins.
- O'Brien, Theresa. (1995). Rhetorical structure analysis and the case of the inaccurate, incoherent source-hopper. *Applied Linguistics*, 16 (4), 442-482.
- Poesio, Massimo and Barbara Di Eugenio. (2001). Discourse structure and anaphoric accessibility. *Proceedings of ESSLLI 2001 Workshop on Information Structure and Discourse Semantics*. Helsinki, Finland.
- Poesio, Massimo, Barbara Di Eugenio and Gerard Keohane. (2002). *Discourse Structure and Anaphora: An Empirical Study* (NLE Technical Note No. TN-02-02): University of Essex.
- Poesio, Massimo, Rosemary Stevenson, Barbara Di Eugenio and Janet Hitzeman. (2004). Centering: A parametric theory and its instantiations. *Computational Linguistics*, 30 (3), 309-363.
- Polanyi, Livia. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12, 601-638.
- Polanyi, Livia. (2001). The linguistic structure of discourse. In D. Schiffrin, D. Tannen and H. E. Hamilton (Eds.), *The Handbook of Discourse Analysis* (pp. 265-281). Malden, Mass: Blackwell.
- Ramachandran, Vilayanur S. and Stuart M. Anstis. (1986). The perception of apparent motion. *Scientific American* (October), 102-108.
- Rambow, Owen (Ed.). (1993). *Proceedings of the ACL Workshop on Intentionality and Structure in Discourse Relations*. Columbus, Ohio: ACL.
- Redeker, Gisela. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14, 367-381.
- Redeker, Gisela. (1991). Review article: Linguistic markers of linguistic structure. *Linguistics*, 29 (6), 1139-1172.
- Redeker, Gisela. (2000). Coherence and structure in text and discourse. In W. Black and H. Bunt (Eds.), *Abduction, Belief, and Context in Dialogue. Studies in Computational Pragmatics* (pp. 233-263). Amsterdam and Philadelphia: John Benjamins.
- Reed, Chris and Derek Long. (1997). *Generating Punctuation in Written Documents* (Technical Report No. RN/97/157). London, UK: Department of Computer Science, University College.

- Reitter, David and Manfred Stede. (2003). Step by step: Underspecified markup in incremental rhetorical analysis. *Proceedings of EACL 4th International Workshop on Interpreted Corpora*. Budapest, Hungary.
- Rösner, Dietmar. (1993). Intentions, rhetoric, or discourse relations? A case from multilingual document generation. *Proceedings of Workshop on Intentionality and Structure in Discourse Relations, ACL* (pp. 106-109). Ohio State University.
- Rösner, Dietmar and Manfred Stede. (1992). Customizing RST for the automatic production of technical manuals. In R. Dale, E. Hovy, D. Rösner and O. Stock (Eds.), *Aspects of Automated Language Generation* (pp. 119-214). Berlin: Springer.
- Rossari, Corinne. (2001). Les relations de discours: Approches rhétoriques, approches pragmatiques et approches sémantiques. *Verbum*, XXIII (1), 59-72.
- Roulet, Eddy. (1995). Geneva school. In J. Verschueren, J.-O. Östman and J. Blommaert (Eds.), *Handbook of Pragmatics* (pp. 319-323). Amsterdam and Philadelphia: John Benjamins.
- Roulet, Eddy, Antoine Auchlin, Jacques Moeschler, Christian Rubattel and Marianne Schelling. (1985). *L'articulation du discours en français contemporain*. Bern: Peter Lang.
- Sanders, Ted. (1986). *De invloed van globale teksteigenschappen 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, Ted and Leo Noordman. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, 29 (1), 37-60.
- Sanders, Ted and Wilbert Spooren. (1999). Communicative intentions and coherence relations. In W. Bublitz, U. Lenk and E. Ventola (Eds.), *Coherence in Spoken and Written Discourse: How to Create It and How to Describe It* (pp. 235-250). Amsterdam and Philadelphia: John Benjamins.
- Sanders, Ted, Wilbert Spooren and Leo Noordman. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15, 1-35.
- Sanders, Ted, Wilbert Spooren and Leo Noordman. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, 4 (2), 93-133.
- Sanders, Ted and Carel van Wijk. (1996). PISA: A procedure for analyzing the structure of explanatory texts. *Text*, 16 (1), 91-132.
- Schauer, Holger. (2000). From elementary discourse units to complex ones. *Proceedings of 1st SIGdial Workshop on Discourse and Dialogue* (pp. 46-55). Hong Kong, People's Republic of China.
- Schauer, Holger and Udo Hahn. (2001). Anaphoric cues for coherence relations. *Proceedings of International Euroconference 'Recent Advances in Natural Language Processing', RANLP 2001* (pp. 228-234). Tzigov Chark, Bulgaria.
- Schiffrin, Deborah. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Schiffrin, Deborah. (2001). Discourse markers: Language, meaning and context. In D. Schiffrin, D. Tannen and H. E. Hamilton (Eds.), *The Handbook of Discourse Analysis* (pp. 54-75). Malden, MA: Blackwell.
- Schilder, Frank. (2002). Robust discourse parsing via discourse markers, topicality and position. *Natural Language Engineering*, 8 (2/3), 235-255.

- Scott, Donia and Clarisse Sieckenius de Souza. (1990). Getting the message across in RST-based text generation. In R. Dale, C. Mellish and M. Zock (Eds.), *Current Research in Natural Language Generation* (pp. 47-73). London: Academic Press.
- Sibun, Penelope. (1992). Generating text without trees. *Computational Intelligence*, 8 (1), 102-122.
- Siegel, Sidney and N.J. Castellan. (1988). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.
- Sparck-Jones, Karen. (1995). Discourse modelling for automatic summarising. In E. Hajicová, M. Cervenka, O. Leška and P. Sgall (Eds.), *Prague Linguistic Circle Papers (Travaux du cercle linguistique de Prague nouvelle série)* (Vol. 1, pp. 201-227). Amsterdam and Philadelphia: John Benjamins.
- Sperber, Dan and Deirdre Wilson. (1995). *Relevance: Communication and Cognition* (2nd ed.). Oxford: Blackwell.
- Stede, Manfred. (2004). The Potsdam commentary corpus. *Proceedings of the Workshop on Discourse Annotation, 42nd Meeting of the Association for Computational Linguistics*. Barcelona, Spain.
- Stevenson, Rosemary, Alistair Knott, Jon Oberlander and Sharon McDonald. (2000). Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, 15, 225-262.
- Stolke, Andreas, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema and Marie Meteer. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26 (3), 339-373.
- Taboada, Maite. (2001). *Collaborating through Talk: The Interactive Construction of Task-Oriented Dialogue in English and Spanish*. Unpublished Ph.D. dissertation, Universidad Complutense, Madrid.
- Taboada, Maite. (2004a). *Building Coherence and Cohesion: Task-Oriented Dialogue in English and Spanish*. Amsterdam and Philadelphia: John Benjamins.
- Taboada, Maite. (2004b). Rhetorical relations in dialogue: A contrastive study. In C. L. Moder and A. Martinovic-Zic (Eds.), *Discourse across Languages and Cultures* (pp. 75-97). Amsterdam and Philadelphia: John Benjamins.
- Taboada, Maite. (2005). Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics* (to appear).
- Taboada, Maite and William C. Mann. (to appear). Applications of Rhetorical Structure Theory. *Discourse Studies*.
- Taylor, P. J. (2002). A partial order scalogram analysis of communication behavior in crisis negotiation with the prediction of outcome. *International Journal of Conflict Management*, 13 (1), 4-37.
- Teufel, Simone, Jean Carletta and Marc Moens. (1999). An annotation scheme for discourse-level argumentation in research articles. *Proceedings of 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL'99)* (pp. 110-117). Bergen, Norway.

- Thomas, Stephen F. (1995). Rhetorical Structure Theory. *Studies in Machine Translation and Natural Language Processing*, 9, 159-174.
- Traxler, Matthew J., Anthony J. Sanford, Joy P. Ake and Linda M. Moxey. (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23 (1), 88-101.
- van Dijk, Teun A. (1977). *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.
- van Dijk, Teun A. (1979). Pragmatic connectives. *Journal of Pragmatics*, 3, 447-456.
- van Dijk, Teun A. (1980). *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse Interaction and Cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- van Dijk, Teun A. and Walter Kintsch. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
- Vander Linden, Keith, Susanna Cumming and James H. Martin. (1992). Using system networks to build rhetorical structures. In R. Dale, E. Hovy, D. Rösner and O. Stock (Eds.), *Aspects of Automated Natural Language Generation* (pp. 183-198). Berlin: Springer.
- Verhagen, Arie. (2001). Subordination and discourse segmentation revisited, or: Why matrix clauses may be more dependent than complements. In T. Sanders, J. Schilperoord and W. Spooren (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects* (pp. 337-357). Amsterdam and Philadelphia: John Benjamins.
- Wahlster, Wolfgang, Elisabeth André, Winifred Graf and Thomas Rist. (1991). Designing illustrated texts: How language production is influenced by graphics generation. *Proceedings of 5th Conference of the European Chapter of the Association for Computational Linguistics (EACL'91)* (pp. 8-14). Berlin, Germany.
- Webber, Bonnie. (2004). D-LTAG: extending lexicalized TAG to discourse. *Cognitive Science*, 28 (5), 751-779.
- Webber, Bonnie, Alistair Knott, Matthew Stone and Aravind K. Joshi. (1999). Discourse relations: A structural and presuppositional account using lexicalised TAG. *Proceedings of 37th Annual Meeting of the Association for Computational Linguistics* (pp. 41-48). College Park, Maryland.
- Webber, Bonnie, Matthew Stone, Aravind K. Joshi and Alistair Knott. (2003). Anaphora and discourse structure. *Computational Linguistics*, 29 (4), 545-587.
- Wiebe, Janyce. (1993). Issues in linguistic segmentation. *Proceedings of ACL SIG Workshop on Intentionality and Structure in Discourse Relations* (pp. 148-151). Columbus, Ohio.
- Wolf, Florian and Edward Gibson. (2004a). Representing discourse coherence: A corpus-based analysis. *Proceedings of the 20th International Conference on Computational Linguistics (COLING)*. Geneva, Switzerland.
- Wolf, Florian and Edward Gibson. (2004b). *A response to Marcu (2003). Discourse structure: trees or graphs?* Unpublished manuscript, MIT, Cambridge, MA.
- Wolf, Florian and Edward Gibson. (2005). Representing discourse coherence: A corpus-based analysis. *Computational Linguistics*, 31 (2), 249-287.
- Wolf, Florian, Edward Gibson and Timothy Desmet. (2004). Coherence and pronoun resolution. *Language and Cognitive Processes*, 19 (6), 665-675.

- Wolf, Florian, Edward Gibson, Amy Fisher and Meredith Knight. (2003). *A procedure for collecting a database of texts annotated with coherence relations*. Unpublished manuscript, MIT, Cambridge, MA.
- Wolf, Florian, Edward Gibson, Amy Fisher and Meredith Knight. (2005). The Discourse GraphBank: A Database of Texts Annotated with Coherence Relations [Corpus]. Philadelphia: Linguistic Data Consortium.