Chapter 2

Coherence Relations: A Survey of Research

2.1 Introduction

This chapter reviews the existing literature on coherence relations, placing it within the wider context of research into discourse coherence. A roughly historical perspective will be taken in presenting the different theories, beginning from the early work in text linguistics and finishing with the latest computational research.

Sections 2.2 and 2.3 provide some preliminaries to the review. In Section 2.2, a working definition of coherence relations is given, broad enough to encompass all the theories to be discussed here. Section 2.3 considers how (if at all) a purely relational theory must be supplemented in order to achieve a complete account of discourse coherence.

In Section 2.4, the many relational theories in the literature are presented. Recently, some of these theories have found useful application in computational domains, particularly in natural language generation: Section 2.5 describes how coherence relations have been adapted to this task, and surveys the different systems which have resulted.

Section 2.6 outlines the central problem to be addressed in this thesis: the proliferation of sets of relations, both in theories of coherence and in generation systems. The lack of a well-established, standard set of relations is already recognised in the literature: each theorist—and each generation system—uses a differently-defined set of relations. This makes it hard to draw comparisons between the various theories and systems; and worse, threatens to undermine the empirical content of the relation construct in general.

2.2 A Working Definition of Coherence Relations

As was noted at the very beginning, the point of departure for a discourse theorist is that a coherent text is more than just a concatenation of sentences. The properties of any given sentence in a text place certain constraints on those of its neighbouring sentences—in other words, there are constraints on the *relations* between the sentences in a coherent text.

However, to say at this point that 'a text is coherent because of the relations between its sentences' is not to make any particular theoretical claim; we are simply re-iterating the phenomenon to be explained. There are as many 'relations between sentences' as there are pairs of possible sentences—clearly what matters for a coherent text is that the relations be of the right kind.

A proper theory of coherence relations thus needs to propose a particular set of relations, and claim that texts are only coherent if they are composed of relations from within this set. At this point, the term 'relation' acquires a new, technical meaning, referring to a class of well-defined theoretical constructs.

Different theories have taken different approaches to the definition of relations. However, there are some concepts which we will take as central to all relational theories: these are outlined below.

2.2.1 A View to Explaining the Notion of Coherence

Relations have been used for a variety of purposes in the literature: for planning the structure of text, for working out the temporal relations between clauses or sentences, for identifying or generating the referents of anaphora, and so on (see Sections 2.4 and 2.5 for details of some of these). However, in this thesis it will be taken as central to a theory of relations that it is concerned at least in part with providing an explanatory account of text coherence. In practice, this conception of relations often underlies accounts developed for specific purposes. (For instance, if a theory of relations is useful in predicting the pattern of anaphora in a text, it is reasonable to assume that this is because it captures something about the ways a coherent text can be structured.) This concern will be a unifying thread behind all the accounts to be considered here.

2.2.2 Text Spans

A theory of relations must specify the units between which relations are required to hold in a text. In all the theories to be considered here, we can begin by thinking of relations as holding between **text spans**—units of text the size of a clause or bigger. In fact, theories often propose that relations hold between more abstract entities, such as propositions or intentions, which are *expressed* in text spans; however, the link between these entities and the spans which express them is sufficient at least to begin to make comparisons. In particular, a common graphical representation is possible for all coherence relations, as illustrated in Figure 2.1.[1] In this diagram, the horizontal lines represent text spans, and the curved line represents the relation between them. The line is labelled with the name of the relation (in this case, EVIDENCE,[2] a relation taken from Rhetorical Structure Theory—see Section 2.4.3 below).

[1] This representation is adapted from that used in Mann and Thompson (1988).

[2] The names of relations will appear in SMALL CAPITALS throughout the thesis.

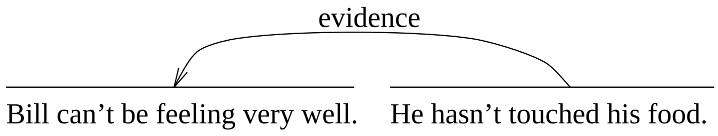

Figure 2.1: Graphical Representation of a Coherence Relation

Text spans can be of different sizes. In some theories, spans are always associated with individual clauses or sentences. In others, they can vary in size from single clauses to whole paragraphs and even larger sections of text. Still other theories allow relations to hold between elements within a single clause. In this thesis, we shall be mainly concerned with spans ranging in size from clauses to paragraphs. It is taken to be a defining feature of coherence relations that they can apply between text spans of any size within this range.

Insensitivity to span size is a useful feature of coherence relations. It permits a hierarchical, recursive account of text, in which small spans are joined together by relations to form larger spans, which are themselves be joined together by other relations. An example is given in Figure 2.2.

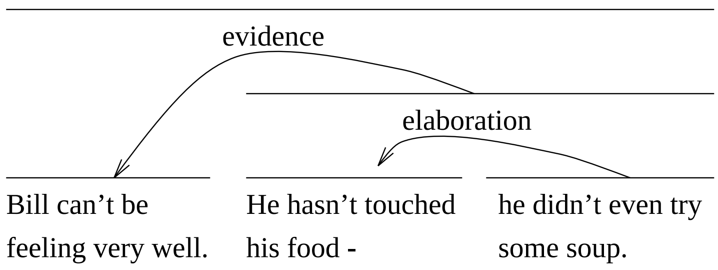

Figure 2.2: A Hierarchical Structure of Relations

2.2.3 Span Structure

As well as a providing a conception of text spans, a relational theory must specify the places where relations are needed in a coherent discourse. Whereabouts in a text must relations be present in order to ensure its coherence? Are they needed between every pair of clauses? Between every pair of paragraphs? How are the various spans in a text supposed to be connected up?

To answer these questions, we need what can be termed a theory of **span structure**, telling us, for any given text, where the coherence relations are expected to be found. Again, there is scope for considerable variation amongst such theories. In some, relations are not expected between every pair of clauses. In others, more than one relation can apply simultaneously between two spans; in others, several consecutive spans can be linked together by a chain of relations; other theories allow relations between discontinuous spans.

2.3 Relations and Focus: Alternative Metaphors for Discourse Coherence

Among other things, the theory of span structure determines how much of the coherence of a discourse is to be attributed to relations. Few, if any, researchers would claim that relations tell us everything we need to know about discourse coherence—they are typically held to account just for some aspects of it. To help delineate the territory to be explained by relations, it is useful to consider some of the other constructs used in theories of coherence; and in particular the concept of **focus**.

Many theories of discourse (eg Sidner (1983), Grosz, Joshi and Weinstein (1983), Brennan *et al* (1987), McCoy and Cheng (1991)) analyse a text in terms of the way its 'focus' moves from one entity to another. At any particular moment while a text is being read, certain objects in its domain of reference are 'in focus': the reader is attending to these elements in particular. The concept of focus has proved extremely useful in modelling anaphora resolution—focused entities are those which can be referred to anaphorically. But it is also used to talk about coherence: the idea is that a writer must obey certain constraints when *changing* the focus from one entity to another. Consider the following two texts:

(2.1) Dow Associates today launched a surprise bid for the beleagured Harris group. Harris reacted speedily, holding an upbeat press conference.

(2.2) *Dow Associates today launched a surprise bid for the beleagured Harris group. Beans, which contain lots of protein, are good for you.

The notion of focus can be used to explain why Text 2.1 is coherent and Text 2.2 is not. At any point in the text, the entities to which the focus can shift are grouped in a list, which Sidner calls the 'potential focus list' and Grosz *et al* call the 'forward center list'. The potential focus list associated with the first sentence in each case contains, among other things, *Harris*, but it does not contain *beans*; Text 2.1 shifts its focus to *Harris* and is thus coherent, but Text 2.2 makes an illegal shift to *beans*, and is thus incoherent.

Clearly, there is a significant possibility for overlap between the metaphors of 'relations' and 'focus'. The coherence of Text 2.1, for instance, might equally be attributed to the existence of a relation (perhaps called something like RESULT or REACTION) between its two sentences.

At the same time, the overlap between the two metaphors is not complete. For instance, the binary nature of relations makes them better suited to an analysis of local constraints on coherence, such as those which might apply between two neighbouring spans regardless of their context. Constraints operative over more than two spans (for example, prohibitions on 'straying from the original topic', or on 'returning to a topic previously closed') might be better dealt with using the focus metaphor. For another thing, relations (as we are thinking of them) do not hold between entities referred to in a text; rather, they link sections of the text itself (or at least the propositions or intentions which underlie it). 'Being in focus', on the other hand, is something more easily attributed to entities being discussed in a text than to portions of the text itself.

Of course, the aspects of coherence targetted by relations and by focus can vary considerably from theory to theory—there is a large degree of flexibility as to how the work should be divided up. The important thing is to make sure that between the two accounts, all aspects of text coherence are covered, and (equally importantly) that there is no redundancy between them. For this reason, when developing a theory of relations it is important to specify exactly which aspects of coherence it is intended to cover. Ideally, what is needed is a clear way of distinguishing the contributions to coherence made by focus and by relations, in such a way as to make best use of both metaphors.

This thesis is primarily concerned with developing a theory of coherence relations; the issue of focus will take a back seat in most of what follows. However, it will reappear with an important role in Chapter 6, when the new theory of relations is to be evaluated, and for this reason it is mentioned here.

2.4 Current Theories of Coherence Relations

To sum up so far: the requirements for a relational theory of coherence are (1) a set of relation definitions; (2) a conception of text spans; (3) a theory of span structure; and (4) some kind of delineation of the phenomena to be accounted for using the relational metaphor. The present section discusses the important relational theories in the literature in the light of these requirements, and examines the differences between them.

I will not go into the details of individual relation definitions in this review, although in many cases a rough idea of the nature of an individual relation is given by its name. My main purpose is to point out the diversity that exists between the many sets of relations that have been put forward.

2.4.1 Some Early Relational Accounts

Attempts to delineate the set of relations that can hold between sections of a discourse date back at least to the 1970s; perhaps the first of note is that of Ballard, Conrad and Longacre (1971). The aim in this study is to catalogue the 'deep' relations which underlie the 'surface' syntactic relations between clauses in complex sentences. Ballard *et al* note that there is no straightforward one-to-one mapping between surface conjunctions and the semantic relationships between the clauses they link—for instance, the conjunction *because* can be used to signal a relation of EFFICIENT CAUSE (*He did it because she wanted him to*) or of GENERIC-SPECIFIC (*They don't taste good because they are bitter*). Conversely, a single semantic relationship can be encoded in different ways: for instance, the sentences *I left before Mary came back* and *I left at about the same time as Mary came back* can be used to describe exactly the same situation, emphasising different aspects of it. The precise roles of deep and surface relations are not completely clear in this account; however, a distinction between the underlying relations in a text and the way they are signalled is retained in many subsequent theories.

The set of deep relations proposed by Ballard *et al* is reworked and refined in Longacre (1983). The new account of relations begins from first principles: a distinction is made between **predications** (typically expressed in clauses) and the **relations** between predications (typically expressed in more than one clause). While the predicate calculus is suitable for modelling clause-internal structure, relations between clauses are therefore described using a variant of propositional calculus. The operations in the propositional calculus—CONJOINING ($\wedge$), ALTERNATION ($\vee$) and IMPLICATION ($\rightarrow$)—are supplemented by a group of TEMPORAL relations to give what Longacre calls the set of BASIC relations; in addition to these he lists a set of ELABORATIVE relations, comprising PARAPHRASE, ILLUSTRATION, DEIXIS and ATTRIBUTION.[3] All of these relations are further subdivided; for instance TEMPORAL relations can be either of OVERLAP or of SUCCESSION. In addition, a parallel set of FRUSTRATED relations is given: these are found in texts where a relation which is expected does not occur. (For instance, the text *They set out for Paris, but didn't arrive* contains a relation of FRUSTRATED SUCCESSION, because there is an expectation that their setting out for Paris will be followed by their arrival.)

A diagram showing the main features of the taxonomy is given in Figure 2.3.

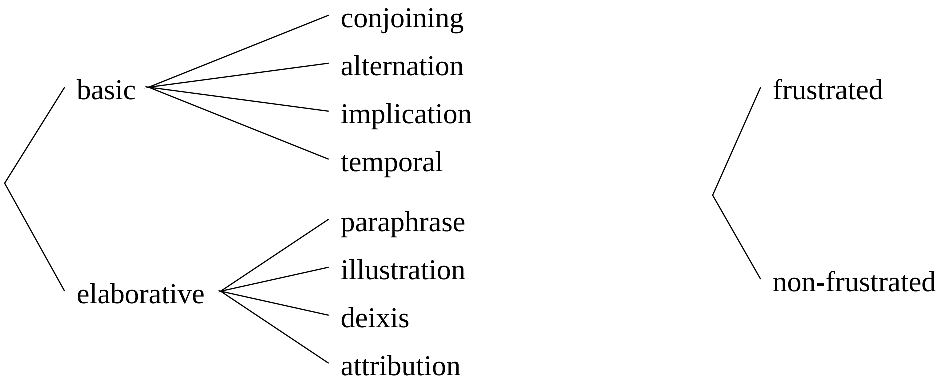

Figure 2.3: The Top Levels of Longacre's Taxonomy of Relations

Grimes' model of relations (Grimes (1975)) gives them a dual role: they provide information, just as clauses do, and they also organise groups of clauses into coherent discourses. Like Longacre, his basic units are clauses (embodying 'lexical predicates') and the relations between them (embodying 'rhetorical predicates'). However, this analysis admits of exceptions. Rhetorical relationships can be found within clauses; for instance, the relation of ALTERNATIVE, typically expressed between two or more clauses, is also found in the single clause *Let's have no more of your neither-here-nor-there observations*.[4] Moreover, the arguments of a lexical predicate can be spread out over several clauses; for instance, the arguments of the predicate 'drop' might be presented as follows: *The stone fell. It hit the ground. Zog made it happen.*).

[3] SMALL CAPITALS will also be used to refer to *groups* of relations identified by a particular theory. Hence, there will be no typographical way of distinguishing between an individual relation and a class of relations.

[4] Though in Grimes' example, it is not very clear what elements in the clause the relation applies between.

The main organising principle in Grimes' taxonomy of relations is a distinction between PARATACTIC and HYPOTACTIC relations. PARATACTIC relations are those whose predicates dominate their arguments in a coordinate fashion; so, for instance, in the ALTERNATIVE relation encoded in *He's in the study or he's in the living room*, the arguments are the two propositions expressed in the two clauses, and the predicate, dominating both propositions, states that exactly one of them is true. HYPOTACTIC predicates relate a central proposition to a subsidiary one; the central proposition then dominates both the subsidiary one and the hypotactic predicate itself. For instance, in the text *He saved the day; he made three touchdowns*, the predicate SPECIFICALLY specifies that the proposition in the second clause gives details about that in the first. This rhetorical proposition, and the specific proposition it refers to, are both subordinate to the proposition in the first clause. NEUTRAL relations are those which can be either PARATACTIC or HYPOTACTIC depending on the context: thus the relation COLLECTION can either group a set of entities together (*We went jogging*), or identify one entity in particular, and associate a secondary group of entities with it (*I went jogging with George and Henry*).

The distinction between PARATACTIC and HYPOTACTIC finds its way in various different guises into many subsequent theories. A second distinction between SYMMETRIC and ASYMMETRIC relations has been less widely adopted—briefly, a relation is SYMMETRIC if the order of its arguments is important and ASYMMETRIC otherwise. It is mainly of relevance to PARATACTIC relations; most HYPOTACTIC relations are ASYMMETRIC.

A representation of Grimes' Taxonomy of Relations is given in Figure 2.4.

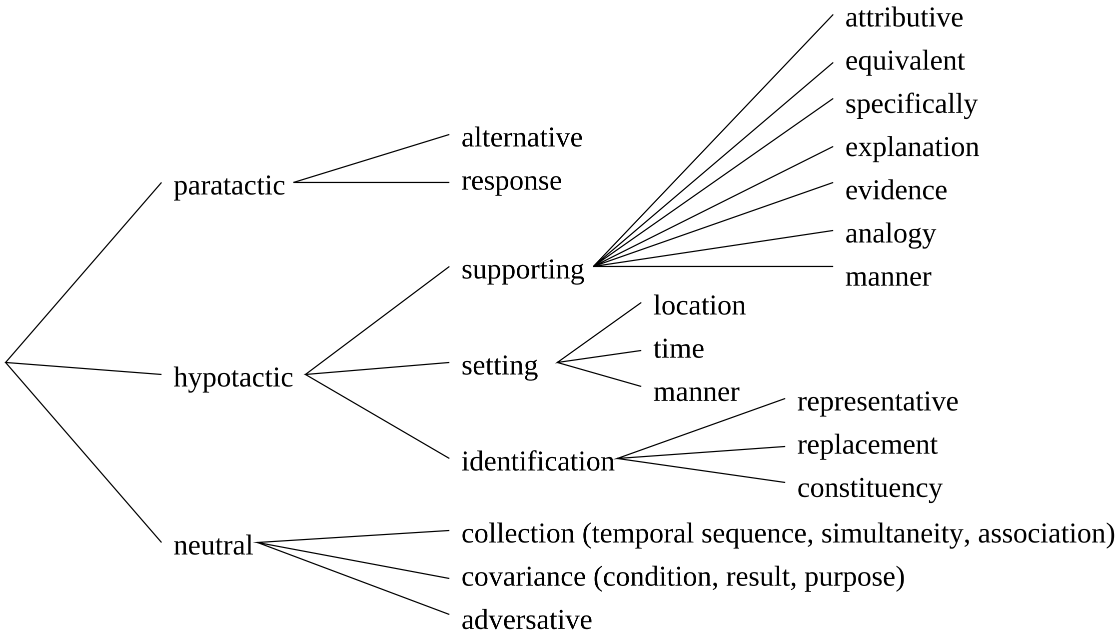

Figure 2.4: Grimes' Taxonomy of Relations

Although their taxonomies are quite different, the theories of Grimes and Longacre have much in common. For one thing, relations are defined independently of surface linguistic phenomena, in terms of the underlying semantics of the clauses which express them—they seem essentially to hold between *propositions*.

A second point of similarity between the two theories is the ability of relations to hold within clauses, as well as between them. Both theories, perhaps because they look beneath the surface structures in text, forfeit a solid conception of the 'text units' between which relations are construed to hold. For instance, Longacre sees an EFFICIENT CAUSE relation in the single clause *Stephen made Kathleen do it*; Grimes sees an ADVERSATIVE relation in the clause *I'd rather have coffee than tea*. It is not clear how, if at all, the semantic units being related in these examples map onto distinct, neighbouring sections of text. It is not even clear that the units being related are whole propositions. In such a situation, a theory of relations starts to encroach on a theory of the semantics of clauses; it becomes hard to see how the work is to be divided up betweeen the two types of theory.

Finally, while both accounts envisage relations applying between large sections of text such as groups of sentences, paragraphs and so on, neither of them has much to say about these higher levels—they both focus on the mechanisms of clause combination within complex sentences. For a fuller account of coherence, relations between sentence complexes and paragraphs must also be investigated.

2.4.2 Cohesive Relations

An alternative conception of relations downplays the importance of the 'deep structure' of text, concentrating on an analysis of the resources available in a language to signal relations explicitly in the surface structure. This approach is adopted by Halliday and Hasan (1976), and also (to some extent) by Martin (1992).

Halliday and Hasan's Conjunctive Relations

Halliday and Hasan (1976) set out to describe the **cohesive** resources in a language; that is, the resources available for linking sentences together in text. Cohesive relations are semantic in nature, and apply both within sentences and between them. However, when found inside sentences, they interact with a theory of sentence **structure** which is outside the scope of Halliday and Hasan's investigation; they are only studied in their 'pure' form, between whole sentences.

There are several types of cohesive relation: **reference**, **substitution**, **ellipsis**, **lexical cohesion**, and **conjunction**. The first four of these can be thought of as dependency structures, where one portion of a text can only be interpreted by reference to something else, either in the preceding or following portions of text or in the environment in which the text is found. Conjunctive relations, on the other hand, are not instructions about how to interpret a portion of text, but a specification of the way in which two already interpretable portions of text are to be linked together. These are the correlates of coherence relations.

Halliday and Hasan's study is primarily of **conjunctive elements** in a language; i.e. its resources for signalling conjunctive relations. Essentially, what is provided is a classification of sentence conjunctions. Several commentators have criticised their reliance on surface features in text for an account of coherence, arguing that texts can be coherent without any explicit signals. However, while a surface analysis certainly dominates in Halliday and Hasan, relations are not tied irrevocably to surface structure:

> it is the underlying semantic relation... that actually has the cohesive power. This explains how it is that we are often prepared to recognise the presence of a relation... even when it is not expressed overtly at all. We are prepared to supply it for ourselves, and thus to assume that there is cohesion even though it has not been explicitly demonstrated.
>
> *Halliday and Hasan (1976), p229*

Having said that, the idea of implicit relations is not discussed at length by Halliday and Hasan, and the mapping between conjunctive elements and 'underlying relations' remains unclear.

The actual classification of conjunctive phrases involves a fourfold distinction between ADDITIVE, ADVERSATIVE, CAUSAL and TEMPORAL relations. The ADVERSATIVE relations cover part of the ground covered by Longacre's FRUSTRATED relations; but here they are not factored out as an orthogonal dimension. An alternative orthogonal distinction is proposed instead, between INTERNAL and EXTERNAL relations. EXTERNAL relations hold between things referred to in a text, while INTERNAL relations hold between elements which are constitutive of the text itself—things such as speech acts. To give a canonical example, consider Texts 2.3 and 2.4:

(2.3) First he switched on the light. Next he inserted the key into the lock.

(2.4) First he was unable to stand upright. Next he was incapable of inserting the key into the lock.

Both texts contain TEMPORAL relations. But in 2.3 the relation is between two events in the world (and hence EXTERNAL); while in 2.4 it is between two *speech acts*, 'I assert that he was unable to stand upright' and 'I assert that he was incapable of inserting the key into the lock'[5] (and is hence INTERNAL). The distinction between INTERNAL and EXTERNAL relations is another one which is echoed in many subsequent relational theories.

The top levels of Halliday and Hasan's taxonomy are given (in slightly simplified form) in Figure 2.5.

By concentrating on an analysis of surface features in text, Halliday and Hasan's theory avoids some of the problems associated with Grimes' and Longacre's accounts. Conjunctive elements are concrete linguistic entities, and it is at least clear whereabouts in a text they are present (even if it is not so easy to decide exactly which units they are linking). But Halliday and Hasan admit that a complete account of text must make reference to unmarked relations. And it is not enough just to recognise the existence of such relations—they must be properly described and individuated.

[5] The interpretation in this case is presumably that the writer is enumerating two points in an argument.

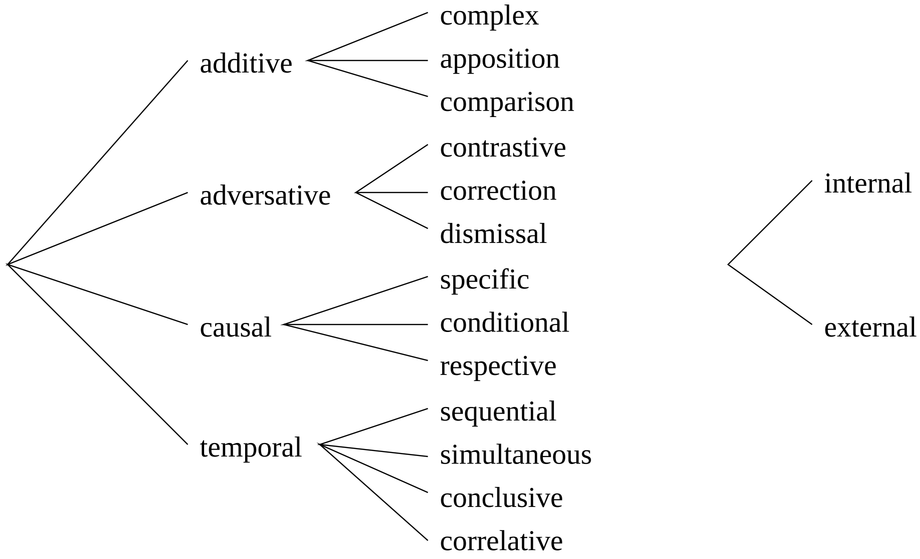

Figure 2.5: The Top Levels of Halliday and Hasan's Taxonomy of Relations

On a separate note, the decision not to consider the relations within complex sentences seems overly cautious. Admittedly, it avoids any possible confusion between a theory of clause semantics and a theory of discourse structure; however, there are so many similarities between clause-combining and sentence-combining phrases that generalisations will surely be lost if only the latter group is studied. Consider the phrases *although* and *however*—the former links clauses, the latter links sentences; but they have very similar functions. Why should they not be accounted for in the same theory?

It should be clear by now that the demands for structural simplicity in a theory and for a complete account of the underlying relations in a text pull in opposite directions. A compromise is needed between 'deep structure' accounts (which tend to fudge the dividing line between theories of clause semantics and discourse structure) and surface accounts (which often leave important features of a text unexplained). It remains to be seen whether a theory can be developed which avoids both of these problems.

Martin's Relations

Martin's (1992) theory of relations follows Halliday and Hasan in its emphasis on explicit linguistic conjunctions—in essence, it provides an alternative taxonomy of connectives. However, it expands on their theory in two respects. Firstly, it gives a better account of the relationship between 'deep' and 'surface' relations: basically, an **implicit relation** exists at a given point in a text if a conjunctive phrase can be inserted at that point.

> As a test for the presence of an implicit connection it can be required that the connection could have been explicit...
>
> *Martin (1992), p184*

This test gives a way not only for identifying the implicit relations in a text, but also for classifying them. The taxonomy of 'underlying relations' will basically mirror that of the devices for signalling them. The test runs into problems with some categories of relations (in particular with the INTERNAL and ADDITIVE relations), but it is nevertheless an improvement on Halliday and Hasan's account.

A second departure from Halliday and Hasan's theory is the decision to look within the sentence for methods of signalling relations. This expands the range of the analysis and enables more generalisations to be made. It also manages to avoid the danger of encroaching on an account of clause-internal semantics; although some clause-internal correlates of conjunctions are discussed (verbs such as *parallelled* and *precluded*, for instance), these are kept quite separate from the class of cohesive conjunctions.

The top levels of Martin's taxonomy of relations are given in Figure 2.6. The tax-

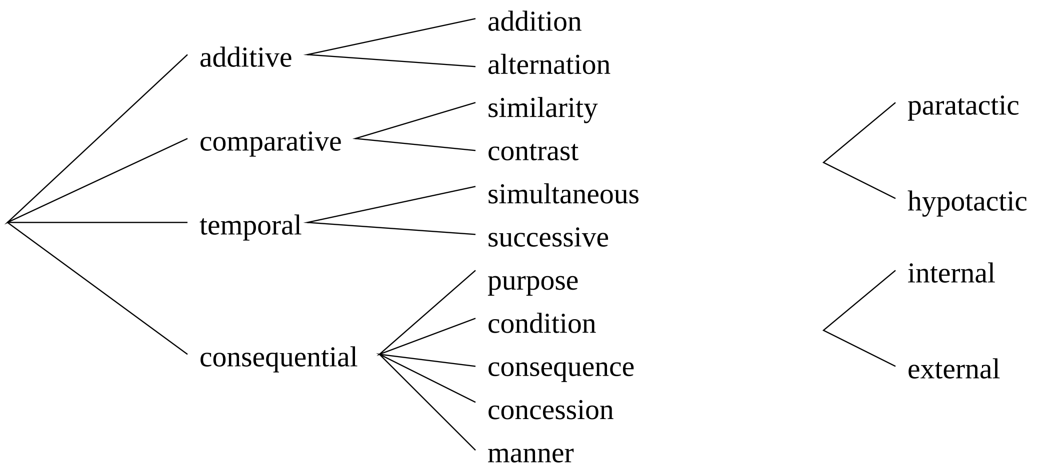

Figure 2.6: The Top Levels of Martin's Taxonomy of Relations

onomy is designed as a **systemic network**, framed in terms of the choices available to the writer. The writer begins by making high-level decisions about the nature of the relation, and these generate progressively lower-level choices until a decision about a particular conjunction can be made. The central division into ADDITIVE, COMPARATIVE, TEMPORAL and CONSEQUENTIAL relations is supplemented by orthogonal distinctions between INTERNAL and EXTERNAL (taken from Halliday and Hasan) and between PARATACTIC and HYPOTACTIC (reminiscent of Grimes's taxonomy). At lower levels, an extremely subtle classification is effected, which distinguishes around 100 different types of conjunction.

2.4.3 Computational Theories of Relations

In the 1980s, computational linguists became interested in coherence relations as potentially useful constructs in discourse processing applications. A number of distinctively 'computational' theories of relations appeared, characterised by the use of A.I. concepts such as knowledge representation, planning and recursion. Three of these theories will be outlined in this section.

Hobbs' Theory

Hobbs' theory (1985) emphasises the amount of world knowledge that is required to interpret discourse. He considers the following text by way of illustration:

(2.5) John took a book from the shelf. He turned to the index.

It is clear to the reader that *the index* referred to in the second clause is that of the book which John has just taken; but to make this inference automatically requires a great deal of knowledge about what people do with books, what indices are, and so on. The relations Hobbs proposes are defined in terms of the different kinds of inferences which the reader needs to draw in order to make sense of a text.

Hobbs identifies four types of inference, and accordingly, four categories of coherence relations. He notes firstly that a discourse can be coherent because it talks about coherent events in the world; events such that if one is known, the other one can be inferred given appropriate background knowledge. Two portions of text describing two such events are said to be linked by an OCCASION relation. This relation subdivides into relations like CAUSE and ENABLEMENT.

Secondly, discourse coherence can be due to the fact that the speaker has some rational structure of goals for producing a discourse. Sometimes these goals are referred to fairly explicitly in the text:

(2.6) Did you bring your car today? *Mine is at the garage.*

Hobbs accounts for such references with the class of EVALUATION relations, which associate some portion of discourse with a piece of 'metatalk' about why that portion is in the discourse.

Thirdly, a discourse will only be coherent if what the speaker says can be linked to what the hearer already knows. This idea provides the basis for the BACKGROUND and EXPLANATION relations, which perform this function.

Fourthly, the hearer of a discourse has only limited processing resources: therefore a coherent discourse makes explicit certain inferences which, given greater resources, he could be expected to make for himself. EXPANSION, the fourth class of coherence relation, links such inferences: two clear examples are PARALLEL and CONTRAST.

The set of relations in Hobbs' theory is given in Figure 2.7.

Hobbs' appreciation of the need for a knowledge-intensive approach to discourse interpretation is characteristic of a computational theory. Also characteristic is a well-articulated account of the compositional nature of relations:

> When two segments of discourse are discovered to be linked by some coherence relation, we can consider the two together as a single segment of

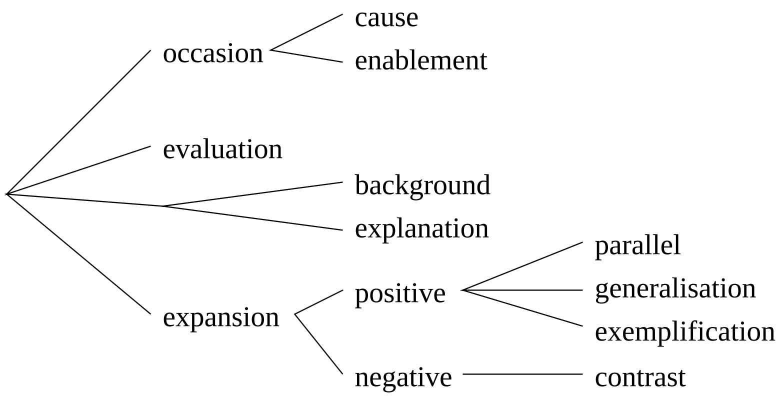

Figure 2.7: The Relations in Hobbs' Theory

> discourse. By recognising coherence relations between segments, we can thus build up recursively a structure for the discourse as a whole.
>
> *Hobbs (1985), p22*

The conception of relations as recursive mechanisms for building up large sections of text has two useful consequences. Firstly, it helps to give substance to the *structural theory* of text called for in Section 2.2.3. This theory, it will be recalled, is required to specify for any text whereabouts relations are expected to be found; Hobbs' suggestion is basically that a tree structure of relations must exist in a text for it to be coherent.

Secondly, the recursive conception of relations suggests for them a *procedural* role in constructing large sections of text. Recursively defined constructs are well suited, for instance, to hierarchical planning formalisms. And thinking about relations in the context of a planning paradigm also suggests how they could be grounded in a *psychological* theory of text processing—at one point Hobbs likens relations to 'text-building strategies', used by the writer to facilitate the task of the reader. This idea which will be taken up in much more detail in the next chapter.

Grosz and Sidner's Relations

Grosz and Sidner's (1986) theory also features recursively defined relations. In this account, **discourse segments** (DSs) are the principal units of structure, and relations hold between these to form larger DSs. However, the primitives used to define relations are different from those of Hobbs: they make reference solely to the **intentions** a writer has in creating a text. Relations actually apply between **discourse segment purposes** (DSPs); an assumption is made that a single overriding intention can be specified for each segment, and it is these intentions which are connected by relations. The fundamental metaphor is of a text embodying the execution of a **plan** pursued by the writer.[6]

[6] Note that although Grosz and Sidner frequently use examples from task-oriented dialogues, they take care in such cases to distinguish the plan required to carry out the task from the plan required to create the text. (See Litman and Allen (1990) for further discussion of discourse plans and domain-level plans.)

Using intentions in relation definitions follows quite naturally from thinking of them in the context of a recursive planning paradigm. Plans are produced to achieve user goals (or in the present case, writer goals); and do so by decomposing a principal goal into a hierarchy of subgoals. What is more, it is easy to see how the same intention can 'underlie' text spans of very different sizes. For instance, the same goal to convince the reader of some fact might require an extended argument for one reader, but a simple sentence for another reader who is already disposed to believe it. Note that it is much harder to claim that two differently-sized texts contain the same 'propositional content'. Intentions therefore seem to provide an ideal way of defining relations which must hold between spans of very different sizes.

There are only two relations in Grosz and Sidner's theory: DOMINANCE and SATISFACTION-PRECEDENCE (see Figure 2.8). Since these are the first intentionally defined relations

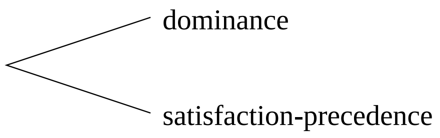

Figure 2.8: The Relations in Grosz and Sidner's Theory

we have come across, it is worth spelling them out in detail:

- DSP2 **dominates** DSP1 (and DSP1 **contributes to** DSP2) if the satisfaction of DSP1 is intended to provide part of the satisfaction of DSP2.
- DSP1 **satisfaction-precedes** DSP2 if both purposes are dominated by some other purpose DSP3, but in order to satisfy DSP3, DSP1 has to be satisfied *before* DSP2.

For example, Text 2.7 contains a DOMINANCE relation:

(2.7) Television is bad for children. They grow up on a steady diet of violence and advertising.

while Text 2.8 (second sentence) contains one of SATISFACTION-PRECEDENCE:

(2.8) Try out the gun by firing off a few rounds. First, release the safety catch; then squeeze the trigger gently.

As well as defining their relations differently, Grosz and Sidner also adopt a slightly different conception of compositionality to that proposed by Hobbs. While Hobbs sees a relation between two adjacent spans as forming a new composite span, Grosz and Sidner's composite discourse segments *include* the segments which they dominate. Thus in Text 2.7, the subordinate span is the second sentence, and the dominant span is the first and second sentences together.

A final attractive feature of Grosz and Sidner's theory is its account of the interaction between relations and focus. Associated with every discourse segment is a **focus space**, and at every point in a text a **focus stack** is given which models the reader's focus of attention as the discourse proceeds (elements at the top of the stack being 'more salient' to the reader than elements lower down). The metaphor of a stack is another import from computational theories. Its pushes and pops are determined by the DOMINANCE relations in the text: if the segment S_2 dominates a sub-segment S_1, then moving into S_1 causes the focus space associated with S_1 to be pushed onto the stack, and leaving S_1 causes it to be popped off the stack.

Grosz and Sidner's theory is the first to look in detail at relations between larger sized units of text; indeed most of their examples are of high-level relations. In fact, the theory does not provide a very complete account of lower level relations such as those between single clauses or sentences. Even 'atomic' discourse segments are often larger than a single sentence. The rationale for these units is based on studies of how 'naive' subjects segment discourses (1975) and of how speech rate correlates with segment boundaries (1975). However, whether or not these studies are convincing, there seems no reason in principle why the analysis of text units should not descend at least to the level of individual sentences.

Rhetorical Structure Theory

We turn now to the third computational theory to be discussed—rhetorical structure theory (RST). This theory, developed mainly by William Mann and Sandra Thompson, is presented in a number of papers; in this thesis I shall for the most part be drawing on the account in Mann and Thompson (1988), which is the most comprehensive.

The central constructs in RST are **rhetorical relations**. Text coherence is attributed principally to the presence of these relations; unlike Grosz and Sidner, Mann and Thompson do not envisage an important role for other constructs such as focus. The claim is that the relations in RST suffice to analyse 'the vast majority' of English texts; exceptions are only made for very unusual texts like poems and legal documents.

Rhetorical relations are defined functionally, in terms of the effect the writer intends to achieve by presenting two text spans side by side. In this respect, they resemble Grosz and Sidner's relations. However, there are also several differences between the two types of relation.

Firstly, RST relations do in fact make some reference to the propositional content of spans, as well as to the intentions of the writer in putting them forward. For instance, the MOTIVATION relation specifies that one of the spans 'presents' an action to be performed by the reader; the SEQUENCE relation specifies that a succession relationship must exist between the related spans. RST relations are in fact defined using five **fields**—only one of these explicitly represents the effect of the relation; the others represent the various different constraints that must be satisfied in order to achieve this effect, and these are specified using a mixture of propositional and intentional language.

Secondly, Mann and Thompson go out of their way to rule out a connection between the set of relations and the linguistic devices used to signal them. This goes beyond the claim that the relations in a text need not be signalled—they further suggest that

> some types of rhetorical relations have *no* corresponding conjunctive signals.
>
> *Mann and Thompson (1988), p45 (my italics)*

In this, their theory differs from Grosz and Sidner's (and many others besides), in which at least an informal link is made between underlying relations and the linguistic devices for marking them.

A third novel feature of RST is its concept of **nuclearity**. As well as representing the relationship between two text spans, rhetorical relations also convey information about which span is more central to the writer's purposes. The **nucleus** is the more central span, and the **satellite** is the less central one.[7] Mann and Thompson contend that the majority of text is structured using nucleus-satellite relations; although some relations—termed **multinuclear**—do not exhibit it. (There are two multinuclear relations: SEQUENCE and CONTRAST.)

The nucleus-satellite distinction is in some ways comparable to the PARATACTIC-HYPOTACTIC distinction of Grimes and others. But while these are expressed in semantic or even syntactic terms, RST's definition is functional, based on the idea that a writer has more important and less important goals when she sets out to create a text. Nucleus-satellite relations are in fact more reminiscent of Grosz and Sidner's class of DOMINANCE relations. But even here there is a difference: in Grosz and Sidner's model it is hard to talk about the purpose of the dominant span being 'more central' to the writer than that of the subordinate span, because the former purpose actually *includes* the latter.

RST provides a set of around 23 rhetorical relations. The numbers vary slightly from paper to paper, but the central core of relations as presented in Mann and Thompson (1988) are given in Figure 2.9. The top-level distinction in this taxonomy is between SUBJECT-MATTER and PRESENTATIONAL relations. SUBJECT-MATTER relations have as their effect that the reader *recognize* the relation in question; while PRESENTATIONAL relations have as their effect to *increase some inclination* in the reader. Thus SEQUENCE is a SUBJECT-MATTER relation (its effect is that the reader recognize that the two related spans present events occurring in sequence) and MOTIVATION is PRESENTATIONAL (its effect is to increase the reader's motivation to perform the action presented in the nucleus span). To some extent, this distinction mirrors Halliday and Hasan's distinction between INTERNAL and EXTERNAL relations. But again, the similarity is far from complete.

Like the other computational theories of relations, RST has a strong structural account of text. It begins with an independent definition of 'text span'—for Mann and Thompson, the size of the atomic units of text analysis is arbitrary, but they should have independent functional integrity. The clause is selected as the minimal unit of organisation; thus text spans are clauses, or larger units composed of clauses. Unlike Grosz and Sidner, relations must hold between non-overlapping text spans. (An ex-

[7] A test for nuclearity is to delete the satellite span of a given relation and see if the resulting text still makes sense. The prediction is that without the nucleus, the significance of the material in the satellite will not be apparent, while the nucleus should be able to stand by itself. This test is rather blunt, and not completely reliable, but it is useful at least in giving some substance to Mann and Thompson's notion of nuclearity.

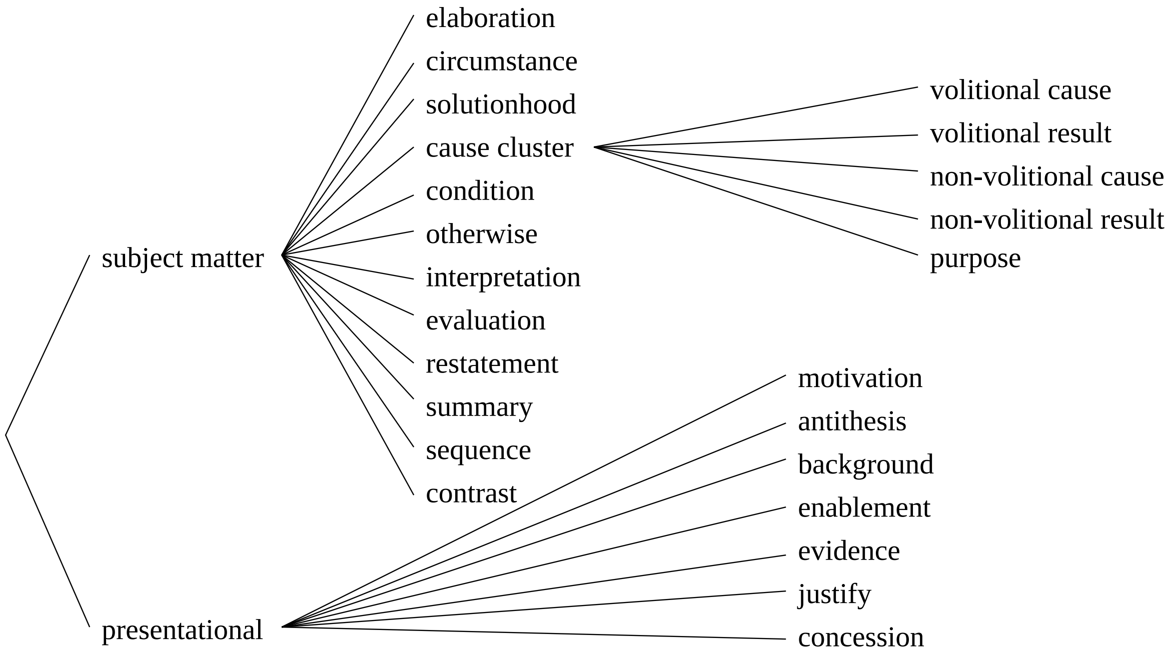

Figure 2.9: Mann and Thompson's Relations

ception to this rule is made for non-restrictive relative clauses: relations are permitted to hold between a matrix clause and a subordinate clause.)

In RST, relations are not mapped directly onto texts; they are fitted onto structures called **schema applications**, and these in turn are fitted to text. Schema applications are derived from simpler structures called **schemas** (see Figure 2.10).

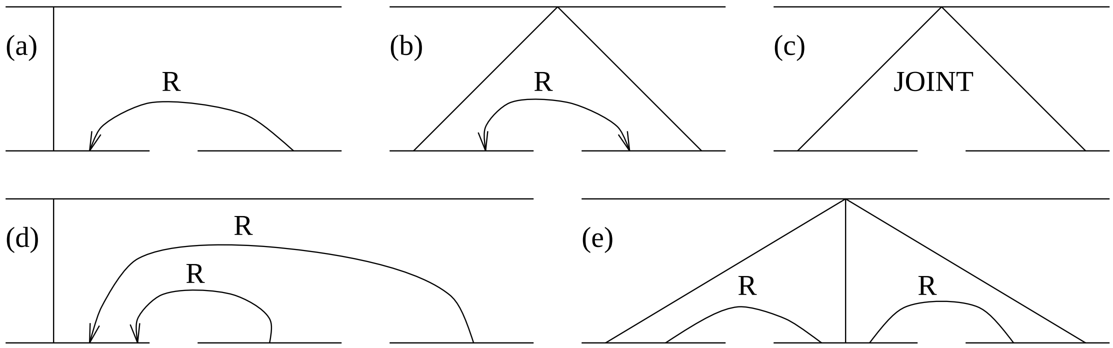

Figure 2.10: The Types of Schema in RST

In this diagram, horizontal lines depict text spans, the labelled lines depict relations between spans, nuclei are picked out by the vertical lines (they are diagonal for multinuclear relations), and all other spans are satellites. From these structures, schema applications are formed, by rearranging the spans in any order and by duplicating spans any number of times. (For the schemas with satellites, only the satellite spans can be duplicated.) Relations are then fitted to the schema applications thus formed—relations which take a nucleus and a satellite are fitted to the single or dual relation schema applications, and the specialised CONTRAST and SEQUENCE relations are fitted to the individual schemas (b) and (e) respectively. The 'joint' schema is for linking pieces of text which are not linked by any RST relations, and is essentially used for representing lists.

A **rhetorical structure tree** is a hierarchical system of schema applications. A schema application links a number of consecutive spans, and creates a complex span which can in turn be linked by a higher level schema application. This enables tree structures to be built—it is a central claim of RST that the structure of every coherent discourse can be described by a single rhetorical structure tree, whose top schema application creates a span encompassing the whole discourse.

An interesting aspect of RST's structural account is its use of the multisatellite schema (illustrated in Figure 2.10 (d)). Note that in this case, there is no relation linking the second and third spans; although they are adjacent, they are only linked indirectly, by virtue of both being related to the first span. (They might, for instance, be two separate JUSTIFICATIONs for a claim expressed in the first span.) Other theorists have chosen to model such structures with a new type of relation; for instance, Halliday and Hasan's INTERNAL SEQUENCE accomplishes exactly this function. The respective merits of the two approaches will be discussed in more detail later in the thesis, in Section 6.2.4.

RST has proved a very influential theory amongst computational linguists, as the next section will attest. Its popularity is perhaps best attributed to a combination of features: the emphasis on a functional conception of relations; the carefully presented set of relation definitions; the simply stated structural theory. It is doubtful whether anyone believes the claims made in RST—but at least it is clearly enough expressed for people to be able to frame their objections to it.

2.5 The Uses of Relations in Automatic Text Generation

In the last few years, coherence relations have begun to feature prominently in studies of natural language generation (NLG). Successive generation workshops (Dale *et al* (1990), McKeown *et al* (1990), Dale *et al* (1992), Zock *et al* (1994), McDonald *et al* (1994)) contain accounts of relations being implemented in generation systems. And two recent workshops, in Maratea (Scott and Hovy (1993)) and Columbus (Rambow (1993)) have been largely given over to a discussion of relations.

The two most popular theories for implementation have been Mann and Thompson's RST and Martin's systemic model. However, as in the theoretical field, no standard set of relations has emerged—the relations in both of these theories have been considerably and variously adapted for their new procedural role. In the following review, I will again be emphasising the differences between the various conceptions of relations which have emerged in the literature.

2.5.1 Hovy's Implementation: Relations as Planning Operators

The original relation-based text structurer was Hovy's (1988) implementation of RST. This system develops the idea of relations as planning operators, defining them in terms of **preconditions** and **postconditions**. Both preconditions and postconditions are expressed in terms of communicative goals—the idea is that a communicative goal can be satisfied in simple cases by producing a single clause, and in other cases by 'applying' an RST relation, whose nucleus and satellite are each characterised in terms of simpler communicative goals. For example, the definition of the operator for the SEQUENCE relation specifies a complex goal as its 'result' (or postcondition), and simpler goals as its 'requirements' (or preconditions):

```
          ___SEQUENCE___
Results:
        ((BMB SPEAKER HEARER (SEQUENCE-OF ?PART ?NEXT)))
Nucleus requirements/subgoals:
        ((BMB SPEAKER HEARER (TOPIC ?PART)))
Satellite requirements/subgoals:
        ((BMB SPEAKER HEARER (TOPIC ?NEXT)))
```

('`BMB SPEAKER HEARER X`' can be read 'achieve the state in which the speaker and the hearer mutually believe that X'. '`TOPIC X`' is satisfied by associating the clause which translates X with the nucleus or satellite of the relation.)

Of course, an operator can only be applied if the appropriate semantic relation holds between its nucleus and satellite. So another part of the precondition for SEQUENCE is as follows:

```
Nucleus+satellite requirements/subgoals:
        ((NEXT-ACTION ?PART ?NEXT))
```

('`NEXT-ACTION`' is a semantic predicate which can be checked by inspection of a database.)

Hovy's structurer takes as input one or more communicative goals, and a set of clause-sized 'input entities' which contain the material to be generated. On the basis of these, a simple rhetorical structure tree is formed, consisting of one relation, one nucleus and one satellite. The tree is then progressively expanded by a process of adjunction: one of its leaf nodes is selected, and replaced by a new relation whose nucleus is the original leaf node and whose satellite is one of the other input entities to be incorporated. The process continues until all the input entities have been used up.

To produce a piece of text, the finished tree is traversed left-to-right, and the leaves are fed to the sentence generator PENMAN (Mann and Matthiessen (1983)). The sentences thus produced are linked using prototypical keywords; for instance, keywords for the SEQUENCE relation include *then* and *next*. The resulting text should in theory satisfy the original communicative goal, include all the material in the input entities, and in addition, be coherent.

Hovy's system is the first to exploit a powerful new technique. Although it is rudimentary, it demonstrates how the power of a hierarchical planning system can be of real use in text structuring problems—the ability of a planner to deal with complex constraints and to consider all possible solutions to a given problem make it well suited to such tasks. And using relations as operators seems to provide just the right amount of flexibility for the planner: if relations do indeed underlie the phenomenon of coherence, then the space of texts to be searched will include all and only those texts which are coherent. However, Hovy's system is only a start; a number of problems can be identified in its operation.

For one thing, in constraining the set of texts produced from a single input, the planner makes use of the notion of **possible growth points**—a notion not apparent in any of the RST literature. The idea is that the nucleus and satellite of a given relation can only be expanded in certain ways; for instance, the nucleus of SEQUENCE can expand into CIRCUMSTANCE, ELABORATION-ATTRIBUTE or PURPOSE and its 'satellite' can expand into ELABORATION-ATTRIBUTE, ELABORATION-DETAILS or another SEQUENCE. There seems no justification for these constraints—texts which violate them can easily be thought up.

In other respects, the generation process appears underconstrained. Constraints to do with the **focus** of the text being built are not addressed, and the texts produced often seem to violate these, switching in an unprincipled way from subject to subject. A system implementing both relational and focus-based constraints is presented by Hovy and McCoy (1989), which resolves some of these problems.

Perhaps the most serious problem with Hovy's system is the way it characterises the text structuring problem in the first place, in terms of a communicative goal and a set of clause-sized input entities. Beginning from these elements begs the question—how are they themselves determined? It is not realistic to think of text structuring happening after the elements of content to be expressed are decided; for instance, we might first decide we need to justify a claim, and only then search for the material to use in the justification.

2.5.2 Moore and Paris: Relations for Guiding Content Selection in Dialogue

The planner developed by Moore and Paris (1989) (see also Moore (1989), Moore and Paris (1993)) addresses the problem of choosing the material to be generated as well as that of deciding how the material should be structured.

> We believe that the tasks of choosing what to say and selecting a strategy for saying it cannot be divided.
>
> *Moore (1989), p67*

Their planner forms part of the interface for an expert system which gives advice in various different domains, and reacts to users' followup questions. It is thus able to draw on the domain-specific information needed to decide on what advice to give.

The technique for interleaving content selection and text structuring tasks centres around the representation of the **intentions** of the text to be produced. For Moore and Paris, it is important that the intention behind every part of the text be represented in the text plan, so that if the plan fails in some respect and the reader has to ask for clarification, it can be determined what went wrong and how to put it right. In Hovy's structurer, the intention underlying the whole text is represented—it is given in the communicative goal specified at input—but the motivation for including additional input entities is not. If some part of the text is unsuccessful and the reader queries it (for instance, by saying 'I don't understand that last point'), what is required is a new attempt to achieve the intended effect of that portion of text. However, if its intended effect is not explicitly represented, this will not be possible: knowing the relation which links the text segment to the rest of the discourse is not sufficient, because frequently what is called for is a new explanatory strategy involving other relations. For instance, if the reader does not understand a concept by use of an ANALOGY, a CONTRAST might be given, although the intention (to explain the concept) is the same in both cases.

Thus it is claimed that there is no simple one-to-one mapping between the relations in a text and the intentions that underlie them. In one sense this is confusing: both Hovy and Moore and Paris work with RST relations, which are defined precisely in terms of the effects they are intended to achieve. The point becomes clearer when it is realised that the intentions behind a span of text can be described at many different levels. Consider Text 2.9:

(2.9) The window's open.

Underlying this text are any (or all) of the following:

- The intention to say "The window is open".
- The intention to convey to the hearer that the window is open.
- The intention to convey to the hearer that the room is cold.
- The intention to make the hearer close the window.
- The intention that the room warm up.

Moore and Paris' point is that the intentions used to define RST relations are often not abstract enough to allow effective backtracking in the case of a local failure.

In a later paper (Moore and Paris (1993)) this point is refined. They claim that for the PRESENTATIONAL relations in RST, the effects specified in the definitions are deemed adequate for generating alternative strategies. (The effects of all these relations, it will be recalled, are to 'increase some inclination in the reader'.) For the SUBJECT-MATTER relations, whose effects are simply that the reader recognise the relation in question, a specification of deeper level intentions is demanded to avoid the backtracking problems outlined above.

Clearly, the intentions used in defining some RST relations are insufficient for some purposes. However, the level of abstraction at which intentions must be specified might be to some extent genre-dependent—in an explanatory dialogue, we can expect PRESENTATIONAL relations to predominate. But in other genres, such as narratives, it is difficult to specify a more abstract goal than the simple representation of content. To give an example: the CAUSE relation can be used as part of an argument, as in Text 2.10, or in a story, as in Text 2.11:

(2.10) Bill must have been absent, because he was sick.

(2.11) Bill was absent from school that day, because he was sick.

The argument in 2.10 hinges on the fact that being sick typically causes absence. But expressing the causal relation is not the writer's primary objective; if the reader is unconvinced, an alternative strategy (perhaps to report that no-one saw him at school) could be attempted. However, if the causal relation is presented as part of a story being told to the reader, as in 2.11, it is hard to identify a goal beyond that of simply relating the cause of Bill's absence. To take another example—the goal behind a SEQUENCE relation in a story is likely to be just to tell the reader what happened next.

2.5.3 Systems using Multiple Levels of Analysis for Relations

A further extension of the notion of intentions is suggested in Moore and Pollack (1992). Here, it is proposed that every discourse should be analysed on two levels, firstly in terms of its intentional (presentational) relations and secondly in terms of its informational (subject-matter) relations. The arguments given are similar to those in the previous section. While the case of purely narrative texts seems to argue against identifying an intentional and an informational relation in every case, there certainly seems some merit in factoring out the two components.

A similar decomposition of relations is proposed by Maier and Hovy (1991) and Maier (1993), and implemented in Hovy *et al* (1992). In these studies, three types of relation are identified, mirroring Halliday and Hasan's (1976) classification of the three functions of language. IDEATIONAL and INTERPERSONAL relations correspond roughly to RST's SUBJECT-MATTER and PRESENTATIONAL relations respectively. TEXTUAL relations serve to 'structure text'—the archetypal examples seem to be the signallers of Halliday and Hasan's INTERNAL SEQUENCE (*firstly*, *secondly*, *moreover* and so on). Maier and Hovy allow that more than one type of relation can hold between two spans in a text; however, there is no need for every type of relation to be present at every point. Thus they allow that descriptive texts are relatively poor in INTERPERSONAL relations.

However, some aspects of this classification make for confusion. For one thing, the class of TEXTUAL relations seems rather different from the other two. It contains not only 'linearising' relations such as those marked by *firstly* etc, but also relations such as 'relnamecomparative', 'relnameconcessive' and 'relnametemporal', which seem to perform much more than a purely ordering function. It is plausible to claim that spans linked by *moreover* are not related by any IDEATIONAL or INTERPERSONAL relations, but merely by the fact that they appear as consecutive items in a text; but these other relations seem quite clearly to convey either propositional or intentional content. There thus seems to be considerable redundancy within the system of relations as a whole.

A second objection to the classification comes from Moore and Pollack (1992): they object to the large number of possible combinations of relations that results from the assumption of three independent relation types. However, this objection seems less telling; Moore and Pollack's use of two independent levels of analysis gives rise to a similar combinatorial problem. Clearly what both theories need to provide is a specification of any contingent *constraints* between the independent components of the analysis. In fact, both theories do provide some initial specifications of such constraints (see Moore and Paris (1993) p670, Maier (1991) pp42–43); but neither of them is yet complete.

2.5.4 Relations in Text Realisation

Attention has also been focused on RST relations as guides to the linguistic marking of text structure.

Although Mann and Thompson strenuously avoid any reference to surface syntactic phenomena in their relation definitions, many correspondences can in fact be found. Several techniques for signalling RST relations are presented by Scott and de Souza (1990), who outline a number of heuristics for guiding the textual realisation process, informed by a combination of stylistics and psycholinguistics. The heuristics, motivated individually, include using embedded relative clauses to signal ELABORATION, and using paratactic and hypotactic coordinators to signal MULTINUCLEAR and NUCLEUS-SATELLITE relations respectively. The utility of such heuristics has given many researchers cause to question Mann and Thompson's policy of downplaying the link between relations and surface linguistic structures. It should certainly not be assumed that any one-to-one mapping between relations and linguistic forms will be found; as Scott and Paris (1995) point out, it is often impossible to predict the structure of a text on the basis of its surface characteristics alone. However, the development and refinement of sophisticated heuristics for marking relations continues to prove useful—see in particular the system proposed by Moser and Moore (1995).

On another front, Rösner and Stede (1992, 1992) and Knott (1991) have suggested ways of making the realisation of a relation sensitive to its hierarchical position in a rhetorical structure tree. Different strategies are appropriate for different levels; of particular interest are high-level marking techniques which involve whole clauses, such as *there are two reasons for this* or *this happened as follows*. At the other end of the spectrum, researchers (notably Vander Linden *et al* (1992)) have investigated ways of signalling RST relations within single clauses. Available techniques include nominalisation and the use of adverbial phrases; for instance the PURPOSE relation can be marked as follows:

(2.12) Follow the steps in the illustration below *for desk installation.*

A final issue of importance is the decision about when to mark a relation and when not to. If relations are always signalled using the most specific available conjunction, the texts which result are often stylistically awkward, making explicit information which is easily inferrable by the reader. Oberlander and Lascarides (1991; see also Lascarides and Oberlander, 1992) introduce the notion of **laconic text** to represent text in which the only relations to be marked are those not inferrable from context and world knowledge. They propose an algorithm for generating such texts using a system of defeasible rules to represent the different types of knowledge that are needed.

2.5.5 Relations in Multilingual Generation Systems

A final use of coherence relations in generation systems is as a language-independent intermediate representation of text structure. Such a representation is particularly useful in multilingual generation systems—relations model the intentions underlying a text, and it seems plausible that these are invariant during translation. A modular approach to multilingual generation is then afforded, whereby it is only in the textual realisation of rhetorical structure trees that different strategies are required for different languages.

Several multilingual systems currently make use of relations in this way; among them Rösner and Stede (1992), Bateman et al (1993). However, there are also studies which call into question the invariance of relations across translations—see for example Delin *et al* (1994). The debate here seems to hinge on the level of abstraction at which relations are intended to represent the information conveyed by a text.

2.6 The Proliferation of Relations, and its Problems

We turn now to the central point to be made in this chapter—that the diversity amongst the many alternative sets of relations, as well as being confusing in its own right, is symptomatic of a deeper confusion about what it is about a text that relations are actually modelling.

I will begin in Section 2.6.1 by looking at the diversity of relations in computational applications. In Section 2.6.2 I will turn to the more serious differences that exist between relational theories, and make some suggestions about what might be responsible for them.

2.6.1 Differences between Generation Systems

All the systems reviewed in Section 2.5 draw principally on RST's set of relations. But despite this, there is a surprising amount of variability between researchers in the sets of relations they use. No two systems use exactly the same set of relations; and no system uses exactly the set of relations proposed in RST.

Departures from RST's original set are of many different types. RSTrelations have been subdivided—for instance, Rösner and Stede's STEP-SEQUENCE is a specialisation of the RST relation SEQUENCE. They have also been amalgamated to form new relations—for instance, Scott and de Souza combine the relations VOLITIONAL-CAUSE, NON-VOLITIONAL-CAUSE, VOLITIONAL-RESULT, NON-VOLITIONAL-RESULT and EVIDENCE, for the purposes of textual realisation. In other cases, relations seem to be defined orthogonally to those in RST: for instance, Hovy et al's (1992) ANALOGY covers some of the ground covered by the RST relations RESTATEMENT, EVALUATION and ELABORATION; but needless to say, these latter relations are not *always* analysable using ANALOGY. Finally, some new relations have appeared which have no obvious connection with RST relations at all—for example, Rösner and Stede define a relation called UNTIL, which is customised for instructional texts where a given action must be performed until a certain condition is met. In short, there are currently hundreds of relations in use—clearly, a significant departure from RST's original set of 23.

This proliferation is partly to be expected: its origin can be traced to a number of features of RST. For one thing, Mann and Thompson are themselves quite flexible about the set of relations in their theory; they are more concerned with establishing the rhetorical relation in general as a useful tool for text analysts.

> Relation definitions have the status of applications of the theory rather than elements of the theory. One might want to change or replace the definitions...such changes are to be expected and do not cross the definitional boundaries of RST.
>
> *Mann, Matthiessen and Thompson (1989), p 48*

The flexibility of RST is in fact held by some to be an advantage of the theory:

> Generally a new domain has dictated modifications to the inventory of relations, but this very adaptability is one of its most useful features.
>
> *Vander Linden et al (1992), p 184*

At the same time, giving theorists such a free hand to choose relations creates some serious problems. For one thing, it makes it hard to compare generation systems amongst each other. But more importantly, it tends to undermine the empirical content of the relation construct in general. The hypothesis that 'virtually any text can be analysed by representing its coherence relations' becomes much less strong if relations can be created whenever they are needed: it is hard to think what evidence could be found which could disprove it. Coherence relations at this point hardly seem to be saying any more than speech act theory; that we must take intentions into account when representing text. The extra claim in RST—that text is coherent by virtue of the relations between its intentions—is virtually unfalsifiable without a method for specifying what is to count as a relation in the first place.

Even incoherent texts can be analysed according to the relations between the intentions in their spans. For instance, the text in (2.13) seems incoherent at first sight:

(2.13) John broke his leg. I like plums.

Yet we could still define a relation which holds between the intentions underlying the spans in this text: perhaps we could call the relation INFORM-ACCIDENT-AND-MENTION-FRUIT. The relation might be multinuclear, with one nucleus intended to convey information about an accident, and the other intended to convey information about fruit. Clearly, we do not want to include these sorts of relations in any principled set of coherence relations.

To place the relational claim on a sounder empirical footing, we need to tighten the constraints on relationhood—for instance, by giving a fixed set of relations, or by giving rules for picking out a set of relations from some larger set of 'potential relations'.

2.6.2 Differences between Relational Theories

A diversity of relations exists not only amongst the various implementations of RST but between RST and the other relational theories. This should be clear from Section 2.4. The number of relations posited in a theory ranges from two to over a hundred; the primitives used for their definition are themselves very diverse; so too are the principles by which relations are organised into taxonomies. Naturally, some of the differences between theories are due to the fact that different theories have different aims—for instance, Grosz and Sidner's theory is geared principally towards modelling the pattern of anaphora in a text, rather than to providing a full set of resources for a text planner. And yet all these theories have, broadly speaking, a common objective—to model the coherence and structure of discourse. They are all apparently trying to explain the same phenomenon; thus it is odd that no consensus about a set of relations is emerging.

One reason for the many different sets of relations is perhaps that none of them seems quite right for expressing all the data. Many theorists acknowledge this themselves: for instance, Grimes (1975) admits that he is 'not completely satisfied with the basis of classification' for his HYPOTACTIC predicates; Martin (1992), in comparing his relations with those of Halliday and Hasan, mentions several areas where 'neither categorisation is completely satisfactory'. The difficulties in finding a suitable classification scheme suggest perhaps that the data to be accounted for is not well captured by a hierarchical taxonomy at all, and that some different formalism might be more appropriate. This idea will be developed in Chapter 5 and beyond.

However, the proliferation of relations also highlights a more fundamental problem, not just to do with the accuracy of relational theories, but again to do with their empirical status. It seems strange, from a theoretical point of view, that we should have a *choice* about which set of relations to use. Theoretical constructs are typically treated as corresponding to *real phenomena* underlying the data they describe, rather than as being purely synthetic: otherwise there seems little point in using the constructs at all. But relational theories tend to downplay this 'realist' conception of relations; consequently, little attention is paid to the question of what it is underlying a text that relations actually model. And yet it is a serious question. 'Coherence' is not just a label applied to an arbitrary group of texts by text analysts—coherent texts are actually produced, and appear in books, newspapers and so on, while incoherent texts (by and large) do not. There must be some reason for this, and if we are looking to relations for an account of coherence, we should expect them to make some reference to it.

In fact, the real situation is not quite as arbitrary as the preceding discussion would imply. There is still some degree of consensus between researchers about which relations to use. (For instance, nearly all of them identify concepts like causality and sequence as important.) But this very fact suggests that intuitions are at work which are not being acknowledged: while researchers rarely attempt justification for their choices beyond an adherence to the needs of descriptive adequacy, their choices are not as diverse as they might be given *only* this criterion.

2.7 Summary

This chapter has reviewed a large number of alternative theories of coherence relations, highlighting the diversity of the different sets of relations that have been proposed. The diversity of relations has of course been noted before—for instance, Hovy (1990) discusses it at length—and it is clear that agreement on a standard set of relations would be a considerable step forward for discourse theorists. Apart from anything else, there is a danger that unless a standard set begins to emerge, the research programme based on coherence relations might grind to a halt. If what one researcher calls a RESULT the next calls a REACTION, and the next an ELABORATION, we will be bound to ask whether there is any point in using such constructs at all.

Clearly what is needed is a standard set of relations. But perhaps more importantly, we need a way of *justifying* a standard set of relations, so that we can begin to argue that one set is more appropriate than another. It is to this latter question that we turn in the next chapter.