

Chapter 6

Discourse

Abstract

Discourse is the area of linguistics concerned with the aspects of language use that go beyond the sentence, and in particular, with the study of **coherence** and **salience**. In this Chapter we discuss proposals concerning two types of coherence—entity coherence (expressed through **anaphora**) and relational coherence— and theories of local and global salience.

Keywords

entity coherence; anaphora; DRT; relational coherence; RST; SDRT; right frontier; salience; local focus; global focus; centering;

6.1 Introduction

The term **discourse** indicates the area of (Computational) Linguistics—also known as **discourse analysis**—that studies the phenomena that are typical of language use beyond the sentence.

One of the key linguistic facts about discourse is that sequences of sentences that do not appear to be connected with each other may be perceived as prob-

lematic (infelicitous, or according to some theories, even ungrammatical) even if each of them is grammatical from the point of view of sentence grammar in isolation (Kintsch and van Dijk, 1978; Zwaan *et al.*, 1995; Graesser *et al.*, 1997) just like 'sentences' consisting of random sequences of words are perceived as problematic even if each of their individual elements is a real word. So whereas the linguistics of sentences is concerned with explicating the factors and rules that make sentences grammatical, one of the main concerns of discourse linguistics are the factors that make discourses **coherent** (Halliday and Hasan, 1976; Kintsch and van Dijk, 1978; Hobbs, 1979; Grosz and Sidner, 1986; Mann and Thompson, 1988; Albrecht and O'Brien, 1993; Gernsbacher and Givon, 1995; Zwaan *et al.*, 1995; Taboada and Mann, 2006).¹

There are a number of reasons why successive sentences may be perceived as being coherent (Zwaan *et al.*, 1995; Knott *et al.*, 2001; Sanders and Spooren, 2001; Poesio *et al.*, 2004; Taboada and Mann, 2006). One reason is because they talk about the same objects (**entity coherence** (Poesio *et al.*, 2004)– also known as **argument overlap** (Kintsch and van Dijk, 1978) or **referential coherence** (Sanders and Spooren, 2001)). Another reason is because they describe events that form a temporal narrative (**temporal coherence**) (Dowty, 1986; Zwaan *et al.*, 1995; Asher and Lascarides, 2003); or are related by causal relations (**informational coherence**) (Hobbs, 1985; Mann and Thompson, 1988; Zwaan *et al.*, 1995; Moser and Moore, 1996b) or, finally, because the sentences are produced to satisfy intentions that are achieved (**intentional coherence**) (Grosz and Sidner, 1986; Mann and Thompson, 1988; Moser and Moore, 1996b). These last three types of coherence are often conflated under the term **relational coherence** (Knott *et al.*, 2001; Sanders and Spooren, 2001; Poesio *et al.*, 2004).

Another factor intensively studied in the linguistics of discourse is **salience** (also known as **topicality**): the degree of importance of an entity or proposition in a discourse (Grosz, 1977; Linde, 1979; Sanford and Garrod, 1981; Gundel *et al.*, 1993; Grosz *et al.*, 1995). Salience has been studied both because it

affects the form of referring expression used to refer to an entity (e.g., pronoun / demonstrative / proper name) (Gundel *et al.*, 1993) and because it affects anaphora resolution (Sidner, 1979).

Perhaps the aspect of discourse that is most studied in linguistics, and certainly in formal semantics, is intersentential **anaphora**. Anaphora is one of the linguistic devices that are used to make text (entity) coherent, but the form of anaphoric expressions is also determined by salience. The other important factor are connectives.

In this Chapter three main topics will be covered: Anaphora and Entity Coherence (Section 6.2),² Relational Coherence (Section 6.3), and Salience (Section 6.4).

6.2 Anaphora and Entity Coherence

The fragment in (6.1) (from the GNOME corpus, (Poesio, 2004)) is a typical example of text that is coherent because its component sentences are **entity continuous** in the sense of the reformulation of Centering theory proposed by (Kibble, 2001): i.e., each sentence refers to at least one entity mentioned in the previous sentence. The first sentence introduces a cupboard which is mentioned again in the second and third sentence. This third sentence introduces Count Branicki, who is referred again in the following sentence, which also mentions the cupboard again.

(6.1) [This monumental corner cupboard]_i follows a drawing by the French architect and ornament designer Nicolas Pineau, who was an early exponent of the Rococo style.

[The cupboard]_i's large scale and exuberant gilt-bronze mounts reflect Eastern European rather than French taste.

[The cabinet]_i was actually made for [a Polish general, Count Jan Klemens Branicki]_j.

An inventory of [Count Branicki]_j's possessions made at [his]_j death de-

scribes both [the corner cupboard]_{*i*} and the objects displayed on [its]_{*i*} shelves: a collection of mounted Chinese porcelain and clocks, some embellished with porcelain flowers.

These subsequent mentions of the same entity are examples of **anaphoric reference**. We begin our survey of linguistic work on discourse by discussing linguistic theories of anaphora, focusing in particular on so-called ‘dynamic’ theories of anaphora, the best-known among which is **Discourse Representation Theory**, or DRT (Kamp and Reyle, 1993).

6.2.1 Context Dependence

The interpretation of many natural language expressions depends on the context of interpretation; in particular, the interpretation of many noun phrases depends on the entities mentioned in the **linguistic context**—the previous utterances and their content. We use here the term **anaphoric references** to indicate expressions that depend on the linguistic context, i.e., on objects explicitly mentioned or objects whose existence can be inferred from what has been said.³ Following the terminology of Discourse Representation Theory (DRT) (Kamp and Reyle, 1993), we will call the set of entities introduced in the discourse situation *U*, for ‘Universe of Discourse’.

6.2.2 Types of anaphoric expressions

As illustrated in (6.1), many types of noun phrases can be used to refer to previously mentioned entities. A particularly clear case are pronouns, whose interpretation entirely depends on the linguistic context. In (6.1) we also find examples of anaphoric definite NPs such as *the cupboard*; indeed also proper names such as the second reference to *Count Branicki* can be dependent on the linguistic context for their interpretation (Wikipedia lists six ‘Count Branicki,’ three of which called ‘Jan Klemens Branicki’). (See also Chapter 27.)

Nominals are not the only expressions whose interpretation is dependent

on the linguistic or visual context in the sense above. Other examples include expressions that could be viewed as the analogous for the verbal interpretation domain of pronouns, such as *pro-verbs* like *did* in (6.2a) and ellipsis such as *gapping* in (6.2b). But just as pronouns are only the most extreme example of context-dependence among nominals, full verbal expressions have a context-dependent component as well. In (6.2c), for instance, the time of listening to the messages is pragmatically determined by the discourse (Partee, 1973; Dowty, 1986; Kamp and Reyle, 1993).

- (6.2) a. Kim is making the same mistakes that I did.
b. Kim brought the wine, and Robin _ the cheese.
c. Kim arrived home. She listened to the messages on her answering machine.

As we will see below, the form of nominals is dictated by the salience of the entities being referred to.

6.2.3 Relation between anchor and antecedent

A text may be considered entity-coherent even in cases in which the semantic relation between the anaphoric expression and its antecedent entity is not of identity—in fact, in the previously mentioned study (Poesio *et al.*, 2004), it was found that about 25% of entity-coherent sentences were so because of a relation other than identity. One example of non-identical anaphoric expressions are cases of **associative** anaphora, in which the context-dependent nominal is related to its anchor by a relation such as part-of, as in (6.3) (also from the GNOME corpus), where *the central door* being referred to in the second sentence is clearly the central door of the cabinet introduced in the first sentence. In these cases, to identify the antecedent a **bridging inference** is generally required (Clark, 1977; Sidner, 1979; Vieira, 1998).

- (6.3) The decoration on this monumental cabinet refers to the French king Louis XIV's military victories.

A panel of marquetry showing the cockerel of France standing triumphant over both the eagle of the Holy Roman Empire and the lion of Spain and the Spanish Netherlands decorates the central door.

6.2.4 Discourse Models

One point that the examples so far should have already made clear is that the universe of discourse U used to identify the anchor Z of a context-dependent referring expression only includes a subset of the objects of a certain type, among which the entities explicitly mentioned in the previous discourse seem especially prominent: for instance, when interpreting *the cupboard* in (6.1), the only cupboard considered seems to be the one mentioned earlier. (This perception is backed up by psychological research suggesting that these examples are not perceived as ambiguous Garnham (2001)). Such considerations are one of the main arguments for the so-called **discourse model** hypothesis (Karttunen, 1976; Webber, 1979; Kamp, 1979, 1981; Sanford and Garrod, 1981; Heim, 1982; Garnham, 1982, 2001) and for **dynamic** models of discourse interpretation. The discourse model hypothesis states that context dependent expressions are interpreted with respect to a discourse model which is built up dynamically while processing a discourse, and which includes the objects that have been mentioned (the universe of discourse U introduced above). This hypothesis may at first sight seem to be vacuous or even circular, stating that context dependent expressions are interpreted with respect to the context in which they are encountered. But in fact two important claims were made in this literature. First, that the context used to interpret utterances is itself continuously updated, and that this **update potential** needs to be modelled as well. Second, that the objects included in the universe of discourse / discourse model are not limited to those explicitly mentioned. The following examples illustrate the fact that a number of objects that can be ‘constructed’ or ‘inferred’ out of the explicitly mentioned objects can also serve as antecedents for context dependent nominals, including sets of objects like the set of John and Mary in (6.4), or propositions

and other abstract objects like the fact that the court does not believe a certain female individual in (6.5). In fact, the implicitly mentioned object may have been introduced in a very indirect way only, as in the case of (6.6), where *the government* clearly refers to the government of Korea, but the country itself has not yet been mentioned either in the text or the title. These implicitly mentioned objects constitute what Grosz (1977) called the ‘**implicit focus**’ of a discourse.

(6.4) John and Mary came to dinner last night. They are a nice couple.

(6.5) We believe her, the court does not, and that resolves the matter. (NY Times, 5/24/ 00, reported by J. Gundel)

(6.6) For the Parks and millions of other young Koreans, the long-cherished dream of home ownership has become a cruel illusion. For the government, it has become a highly volatile political issue. (Poesio and Vieira, 1998)

The idea of discourse model, originally formulated by Karttunen (1976), was then developed by Sanford and Garrod (1981) and Garnham (2001) in psycholinguistics, and made more formal, among others, by Heim (1982) and Kamp (1981) in theoretical linguistics and Webber (1979) in computational linguistics.

The theories developed by Heim and Kamp collectively took the name of Discourse Representation Theory (DRT); DRT has become the best known linguistic theory of the semantics of anaphora, and has served as the basis for the most extensive treatment of anaphora proposed in linguistics, (Kamp and Reyle, 1993), which integrates anaphora within a theory of semantics also covering the other aspects of semantic interpretation discussed in Chapter 5 of this Handbook. In DRT, a discourse model is a pair of a set of discourse referents and a set of conditions (statements) about these discourse referents:

$$\langle x_1..x_n, c_1..c_n \rangle$$

represented in the linear notation of Muskens (1996) as

$$[x_1 \dots x_n | c_1 \dots c_n].$$

For instance, suppose A addresses utterance (6.7a) to B in an empty discourse model⁴. Then according to DRT update algorithms such as those proposed in (Kamp and Reyle, 1993; Muskens, 1996), when we process this utterance, we update the existing discourse model with information contributed by this utterance: that an entity, engine e3, has been mentioned (hence a discourse referent x_1 'representing' that entity gets introduced in the discourse model); and that 'we' (speaker A and addressee B) are supposed to take x_1 . This fact, as well as the fact that x_1 is an engine, are new conditions added to the discourse model. The resulting discourse model is as in (6.7b). Note in particular that interpreting nominal expression *engine E3* has resulted in a new discourse referent being added to the universe of discourse U. (Here and elsewhere we'll ignore illocutionary force and simply treat all such utterances as statements.)

(6.7) a. We're gonna take engine E3

$$\text{b. } [x_1 | x_1 = e_3, \textit{engine}(x_1), \textit{take}(A + B, x_1)]$$

This discourse model is the context in which the interpretation of the following utterance takes place. Say that (6.7a) is followed by (6.8a), which contains a pronoun. This pronoun has only one interpretation in the discourse model in (6.7b)—as having discourse entity x_1 as antecedent. Interpreting utterance (6.8a) —i.e., establishing that an instruction to send engine E3 to Corning—leads to a second update of the discourse model; the resulting model is as in (6.8b) and contains, in addition to the discourse entities and the conditions already present in (6.7b), new discourse entities and new conditions on these entities.

(6.8) a. and shove it to Corning

$$\text{b. } [x_1, x_2, x_3 | x_1 = e_3, x_2 = x_1, x_3 = \textit{corning}, \textit{engine}(x_1), \textit{take}(A + B, x_1), \\ \textit{send}(A + B, x_2, x_3)]$$

Two key contributions of dynamic theories of anaphora developed in formal linguistics have been to show that the construction of such discourse models

can be characterised in a formal way, and that the resulting interpretations can be assigned a semantics just as in the case of interpretations proposed for other semantic phenomena discussed in Chapter 5. The original approach to discourse model construction proposed by Heim (1982) and Kamp (1981) – and later spelled out in painstaking detail by Kamp and Reyle (1993) – was highly idiosyncratic, but later work demonstrated that the methods of syntax-driven meaning composition used in mainstream formal semantics can be used to develop a theory of discourse model construction as well (Heim, 1983; Rooth, 1987; Groenendijk and Stokhof, 1991; Muskens, 1996). In the following Sections of this Chapter, we will see how the basic type of discourse model developed in the dynamic tradition to account for entity coherence can be extended to provide a formal model of other types of coherence, and in particular relational coherence.

6.2.5 Resources for studying anaphora

In recent years a number of corpora annotated with anaphoric information in multiple languages have become available, so that we are now in an excellent position both to study of anaphora from a linguistic perspective and to develop anaphora resolution algorithms (see Chapter 27). Historically, the first such resources to be made available were the MUC and ACE corpora⁵, but the annotation scheme used for these corpora has been much criticised from a linguistic perspective (van Deemter and Kibble, 2000). Resources created more recently, however, follow more standard linguistic guidelines, as well as typically being larger in size. Among these corpora recently made available we will mention first of all the OntoNotes corpora for Arabic, English and Chinese, also annotated with a variety of other semantic information⁶; the ARRAU corpus of English;⁷ the Ancora corpora for Catalan and Spanish;⁸ the Prague Dependency Treebank for Czech;⁹ the COREA corpora for Dutch;¹⁰ the TüBa-D/Z corpus for German;¹¹ the LiveMemories corpus for Italian;¹² and the NAIST corpus for Japanese.¹³

6.3 Relational Coherence

A discourse may also be perceived as coherent if its utterances are temporally or causally coherent (Zwaan *et al.*, 1995), i.e., if it consists of events that can be interpreted as temporally or causally linked (following (Moser and Moore, 1996b), we use the term **informational** coherence to refer to coherence arising from the event structure of the discourse) or if the utterances themselves can be interpreted as forming a coherent argument in that the intentions to express are related (**intentional** coherence) (Grosz and Sidner, 1986; Mann and Thompson, 1988; Moser and Moore, 1996b). These two forms of coherence¹⁴ are normally grouped together under the term **relational** coherence as the linguistic devices most commonly used to indicate these types of connections are connectives—causal connectives such as *because*, temporal connectives such as *after*, etc.

A number of accounts of relational coherence have been proposed in (Computational) Linguistics. The best-known models however, including Rhetorical Structure Theory (RST) and Segmented Discourse Representation Theory (SDRT), are based on assumptions that go back to the work by Grimes (1972) and van Dijk and Kintsch (1983). According to Grimes, a coherent discourse has the structure of a tree (the **content tree**) whose nodes are propositions (henceforth, **discourse units**) and whose arcs express relations between those discourse units (Taboada and Mann, 2006).¹⁵ Grimes called the relations between these units **rhetorical predicates**, but in most subsequent accounts they are called **rhetorical relations** (e.g., (Taboada and Mann, 2006)). Three types of rhetorical relation exist according to Grimes:

hypotactic: these relations relate two or more propositions, one of which is superordinate to the others. For example, the relation evidence relates a proposition staking a claim with one or more additional propositions providing evidence for that claim.

paratactic: these relations relate propositions of equal weight, and therefore represented in the content tree at the same level. An example are connec-

tives like conjunction and disjunction.

neutral: these relations can be either paratactic or hypotactic according to the emphasis (*staging*) used by the authors. (These relations were not particularly studied in following work.)

In Grimes' model, and in all accounts based on this view, the content tree has a recursive structure, in the sense that discourse units include both atomic propositions, i.e., propositions whose predicate is a verb or noun (that Grimes called *lexical propositions*) and propositions whose predicate is a rhetorical relation (that Grimes called *rhetorical propositions*). For instance, in example (6.9), whose structure is illustrated in Figure 6.1, discourse units a., b. and c. are lexical propositions; b. and c. are connected by a paratactic relation (coordination), forming rhetorical relation d.; and a. and d. are related by hypotactic relation of evidence.

- (6.9) a. Parakeets are ideal for people with little room
b. A cage takes very little room,
c. and a small apartment is sufficient space for their flying exercises

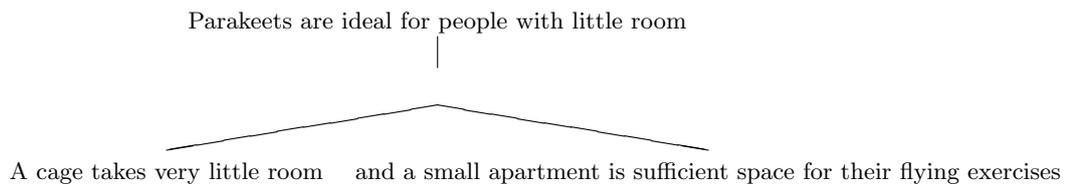


Figure 6.1: Content tree for discourse (6.9)

One of the key claims of the Grimes / RST view of discourse is that there is a connection between coherence and connectivity: the coherence of a discourse depends on all of its units being included in the content tree, fully connected by relations belonging to a repertoire whose specification is a crucial aspect of a theory.¹⁶ Note however that this view requires having relations that capture entity coherence—one example being elaboration, whose ambiguous status is analysed in-depth by Knott *et al.* (2001).

Early empirical evidence for the coherent-discourse-as-a-content-tree view was provided by Meyer (1975)—who showed that the structure of a discourse has effects on recall, with the discourse units higher up in the tree being more likely to be recalled by subjects—and by Kintsch and van Dijk (1978). These and the success of this type of account for discourse analysis led to this view being adopted in a number of accounts, the best known among which is **Rhetorical Structure Theory**, or RST (Taboada and Mann, 2006). We discuss RST in Section 6.3.1. A second very influential strand of research in relational coherence is the work by Grosz and Sidner (1986), who proposed an intentional account of relational coherence in which the discourse units are **discourse intentions** and only two relations between them are assumed. We briefly discuss this proposal in Section 6.3.3 and Grosz and Sidner’s overall framework to discourse in more detail in Section 6.4.2. A third line of research has been concerned with providing an explicit formalization of informational and intentional relations in terms of their inferential properties. The best known work in this direction is that by Hobbs (Hobbs, 1978b, 1979, 1985) and, more recently, by Asher and Lascarides, who integrated an account of relational coherence within the theory of anaphora provided by DRT developing Segmented DRT, or SDRT, which is at the moment the best known linguistic theory of relational coherence (Asher, 1993; Asher and Lascarides, 2003). We discuss the theory in Section 6.3.4.

More modern research has however challenged the assumptions of the content tree view, in at least two directions. In work by Penstein-Rose, Wolf and Gibson, and others, evidence is presented that the relational structure of a discourse is a graph rather than just a tree (Rosé *et al.*, 1995; Wolf and Gibson, 2006). The second novel direction has been to reconsider the way rhetorical relations are defined by grounding more explicitly in linguistic phenomena. Knott (1996) proposed to limit the range of rhetorical relations to those that could be explicitly expressed using a **discourse connective**. Webber *et al.* (2003) expanded this account by proposing to treat discourse adverbials as anaphors. This line of work motivated the creation of the Penn Discourse Treebank,¹⁷ at

present the largest corpus annotated with discourse relations.

6.3.1 Rhetorical Structure Theory and its Variants

Rhetorical Structure Theory is a theory of text organization developed by Mann and Thompson in a long series of papers (Mann and Thompson, 1983, 1988, 1992); an excellent summary of the theory and the issues it raised can be found in (Taboada and Mann, 2006). RST has been very successful both in discourse analysis (e.g., (Fox, 1987)) and in Computational Linguistics, where it has been intensively used especially in Natural Language Generation (e.g., (Hovy, 1993; O'Donnell *et al.*, 2001)) and in Summarization (e.g., (Marcu, 2000)).

RST inherited from Grimes' theory the idea that coherent texts can be characterised in terms of a content tree covering all discourse units, and many of the relations proposed, but introduces a key new idea, the concept of **nuclearity**. This is the hypothesis that in many rhetorical relations—the term **schema** is used in RST—certain units (the **nuclei**) are more important than others.

Every rhetorical relation is represented in RST by a schema which gets instantiated in actual text. A schema indicates how a particular unit of text structure is decomposed into component units, called **spans**. As in Grimes' theory, schemas can be hypotactic or paratactic. The (unpublished) analysis of a short text from the Scientific American in Figure 6.2 illustrates both hypotactic and paratactic schemas.¹⁸

The whole content tree in Figure 6.2 is an instance of the hypotactic Preparation schema, whose **nuclear** component, indicated by the vertical bar, are discourse units 2-5. In this schema, the subordinate unit, or **satellite**, contains material that introduces the reader to the material contained in the nuclear component. This nuclear component is in turn an instance of a hypotactic schema, Background. In this schema, the material in the satellite (2-3) provides background that helps understanding the material in the nucleus (4-5). In turn, the satellite is an instance of the Elaboration schema, whose satellite (unit 3) provides additional information about the entity in the nuclear span (unit 2).

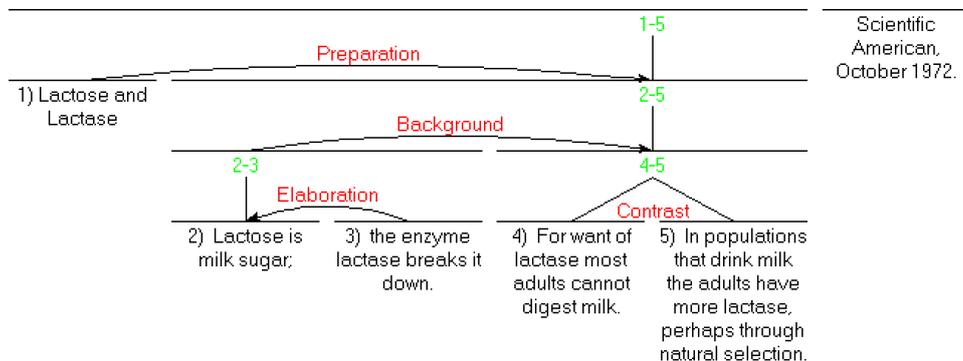


Figure 6.2: An RST Analysis of a short text.

The discourse unit 4-5 is instead an instance of a paratactic schema, Contrast.

A key aspect of the theory is the repertoire of schemas, and much research (by Mann and Thompson and others) over the years has focused on this aspect. Proposals range from smaller sets of about 20 schemas to very large sets of 50 schemas and more (for discussion see (Taboada and Mann, 2006); see also (Hovy and Maier, 1991; Sanders *et al.*, 1992)). Examples of additional schemas discussed in Mann (1984) include:

evidence This schema is analogous to Grimes' **evidence** rhetorical predicate:

the spans related to the nucleus by an **evidence** relation stand as evidence that the conceptual span of the nucleus is correct.

thesis/antithesis This schema is analogous to Grimes' **adversative** rhetorical

predicate, and relates a set of ideas the writer does not identify with and a second collection of ideas that the writer does identify with.

concessive It's not completely clear from the paper what the difference between this schema and the thesis/antithesis one is.

conditional This schema relates a number of satellite clauses which present

a condition, to a nucleus which present a result that occurs under that condition.

inform In this schema, there is a central assertion which constitutes the nu-

cleus, a number of satellites which give an **elaboration** of the nuclear clause, as well as a number of spans which give the **background** for that assertion.

justify In this schema the satellite is attempting to make acceptable the act of expressing the nuclear conceptual schema.

A particularly difficult question is the extent of agreement on a particular RST analysis. Marcu *et al.* carried out extensive analyses of the agreement on different aspects of analysis using RST, finding different degrees of agreement (Marcu *et al.*, 1999). A related issue is the claim, made by Moore and Pollack (1992) on the basis of examples like (6.10), that in some cases discourse units can be related by multiple relations.

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

Moore and Pollack make a distinction between **informational** relations that express a connection between facts and events 'in the world' (such as causal and temporal relations), and / or **intentional** ones that express a discourse intention (such as evidence or concession). (In RST, the terms *subject-matter* and *presentational* relations are used for these two classes of relations (Mann and Thompson, 1988), p. 18.) According to Moore and Pollack, the two units in (6.10) can be viewed as being related both by an intentional evidence relation (with b as a nucleus, and a as a satellite) and by an informational volitional cause one. Furthermore, Moore and Pollack argued that whereas Mann and Thompson claimed that in such cases (which they did observe) one relation had to be chosen, preserving both relations was in fact not only useful to avoid conflicts, but necessary, to account for the flow of inference from both an interpretation and generation point of view. Moore, Moser, and Pollack developed a version of RST called **Relational Discourse Analysis** (RDA) (Moore and Pollack, 1992; Moser and Moore, 1996b) based on these distinctions.

Corpora annotated according to RST and its variants One reason for the decline in popularity of discourse research between 1995 and 2005 was the lack of annotated resources allowing such research to be put on the same empirical footing as other areas of Computational Linguistics. This situation has drastically changed due to the annotation of a number of such resources, and in particular the RST Discourse Treebank (Carlson *et al.*, 2003). A corpus annotated according to RDA also exists, the Sherlock Corpus (Lesgold *et al.*, 1992; Moser and Moore, 1996a).

6.3.2 The Right Frontier

Perhaps the best known linguistic claim rooted in the content tree theory of relational coherence and discourse structure is the so-called **right-frontier constraint**, often associated with the name of Polanyi (Polanyi, 1985) and Webber (Webber, 1991), but also found in other work on the association between discourse structure and anaphora, and which has become a key motivation for SDRT (Asher, 1993; Asher and Lascarides, 2003).

The constraint states that only material introduced in the last discourse unit added to the tree, or in the discourse units which are superordinate to it, are accessible for anaphoric reference or for any other type of discourse-based attachment. Consider the following example from (Lascarides and Asher, 2007). According to Asher and Lascarides, the content tree for sentences (6.11a-e) is as shown in Figure 6.3: b. to e. are subordinate to a. (they stand in a Elaboration relation), and c. and d. stand in a subordinate relation to b. When processing sentence f., the right frontier consists of the last sentence, e., and its superordinate units, a. Because of this, the salmon introduced in embedded discourse unit c. is inaccessible.

(6.11)a. John had a great evening last night.

b. He had a great meal.

c. He ate salmon.

- d. He devoured lots of cheese.
- e. He won a dancing competition.
- f. ??It was a beautiful pink.

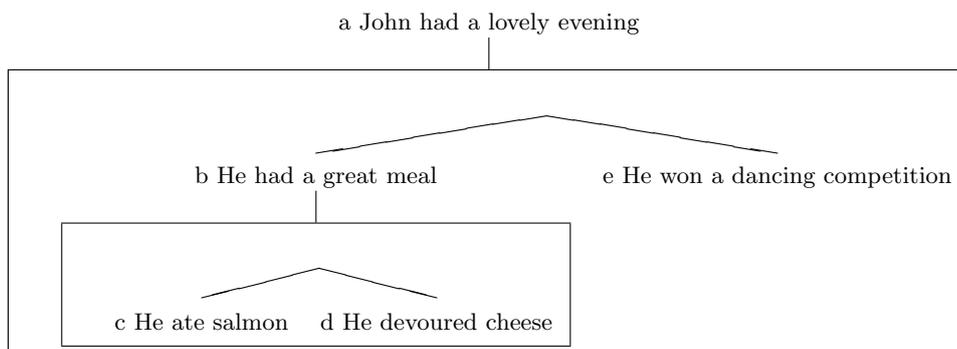


Figure 6.3: Illustration of the Right Frontier Constraint

We will return to the right frontier constraint while discussing discourse structure and anaphora.

6.3.3 Grosz and Sidner’s Intentional Structure

Grosz and Sidner (1986) proposed a radically simplified version of the content tree account of relational coherence. In their account, motivated primarily by dialogue data (that they take as the primary example of discourse), each utterance in a discourse is meant to achieve a (discourse) intention, or **discourse (segment) purpose**; and the text is coherent to the extent that such intentions are connected to form a tree. Examples of typical discourse intentions are “intending that an agent believes some fact” or “intending that some agent believes that some fact supports some other fact”, but can be arbitrarily complicated. On the other end, only two types of relations are assumed to exist between intentions: **dominance** (corresponding to subordination) and **satisfaction-precedence**. In example (6.12), for instance, utterance a. could be viewed as express a discourse segment purpose DSP1 of A to have B carry out an action. Utterances b-d express a new discourse segment purpose DSP2, of B engaging with B to execute the action; this discourse intention is dominated by

DSP1. Then utterances e and following express a discourse intention DSP3 of A to have B carry out the second part of the instruction; DSP3 is also dominated by DSP1, but it is also satisfaction-preceded by DSP2 in that satisfaction of that discourse intention is a prerequisite for the satisfaction of DSP3.

(6.12)a. A: Replace the pump and belt please.

b. B: OK, I found a belt in the back.

c. is that where it should be?

[replaces belt]

d. OK it's done

e. A: Now remove the pump.

f. First you have to remove the flywheel ...

...

As can be observed already from this example, the drastic reduction in the number of relations in the Grosz and Sidner framework is achieved by introducing a notion of 'intention' that is very unspecified, making an intentional analysis of discourses very difficult so that to our knowledge there are not large-scale corpora annotated in this style. A notion of intention derived from Grosz and Sidner's but much more specified was introduced and used for annotation in the already mentioned RDA framework (Moser and Moore, 1996b). A fully formal account of intentions in discourse was introduced in the SDRT framework, discussed next.

6.3.4 Logical Models of Relational Coherence and SDRT

Hobbs' formalization of coherence relations Hobbs developed in (Hobbs, 1978a, 1979, 1985) a theory of coherence based on a formalization of coherence relations in terms of the beliefs and goals of the agents producing and interpreting the discourse.

According to Hobbs, the situation in which a discourse takes place can be

described as follows:

1. The speaker wants to convey some message;
2. The message is in service of some goals;
3. The speaker must relate what is new and unpredictable in the message to what the listener already knows;
4. the speaker should ease the listener's difficulties in understanding.

Hobbs claims that coherence can be characterised by means of a small number of binary **coherence relations** (CRs) between a current utterance and the preceding discourse. The CRs are definable in terms of the operations of an inference system, hence can be inferred by a system able to construct chains of inference and to make similarity comparisons (more on this below). Corresponding to each of the requirements on discourse above, there is a class of coherence relations that help the speaker satisfy that requirement.

For example, suppose the system is given the coherent text

(6.13) John can open Bill's safe. He knows the combination.

The system has definitions of the coherence relations in terms of the propositional content of the sentences in the text. For example, it has a definition of the coherence relation Elaboration which goes as follows:

Elaboration: A segment of discourse S1 is an **Elaboration** of segment S0 if the same proposition P can be inferred from both S0 and S1, and one of the arguments of P is more fully specified in S1 than in S0.

In order to recognize that an elaboration relation exists between the two sentences in (6.13), the system (1) has to know that if X knows the combination of a safe then X can open it, therefore, the same proposition

$can(J, open(Safe))$

can be inferred from both sentences and (2) has to recognize that this situation matches the definition of the **Elaboration** coherence relation above.

Hobbs also developed a view of the process by which the process is constructed based on abductive reasoning, a particular form of defeasible reasoning (Hobbs *et al.*, 1993).

Hobbs' inferential theory of coherence inspired work on SDRT (below) as well as work by Kehler (Kehler, 2002; Kehler *et al.*, 2008).

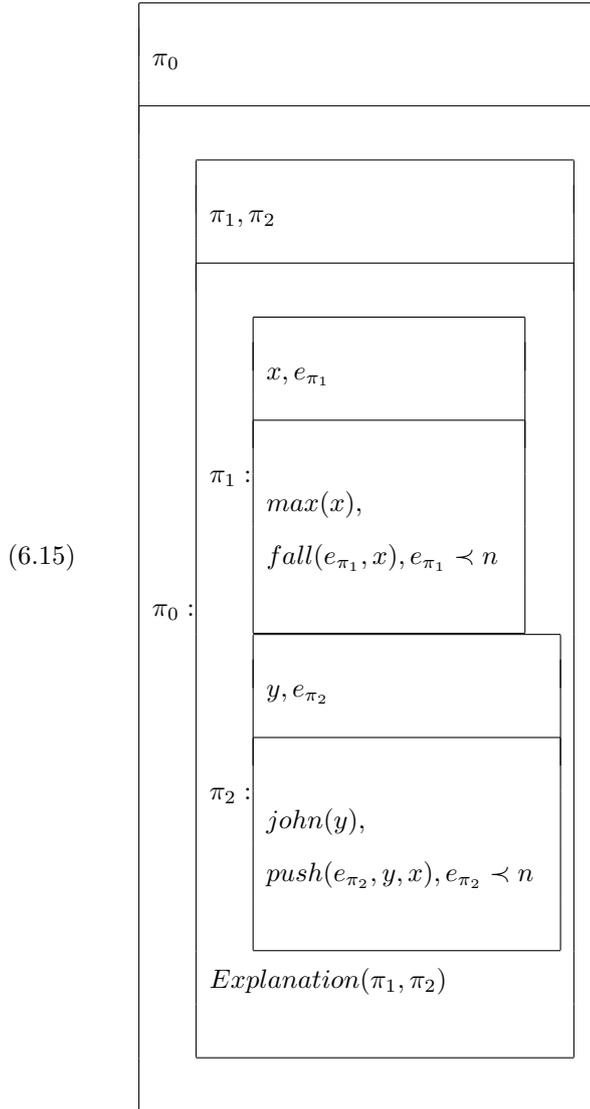
Segmented Discourse Representation Theory (SDRT) SDRT (Asher, 1993; Asher and Lascarides, 2003; Lascarides and Asher, 2007) is an extension of the 'dynamic' view of semantics and discourse modelling exemplified by DRT and discussed in Section 6.2. Asher and Lascarides had two main aims in developing their theory: incorporating in the dynamic view the notion of coherence and the pragmatic constraints on interpretation deriving from the content tree view of relational coherence; and developing a formulation of the process by which the discourse model is constructed consistent with current thinking according to which such process is a form of defeasible reasoning, as proposed by Hobbs *et al.* (1993) and by now assumed by most computational linguists. Much of the theory, in particular as exposed in (Asher and Lascarides, 2003), is concerned with developing this view of discourse processing as defeasible reasoning while taking into account complexity considerations. However, we will only discuss here the theory's formalization of relational coherence.

The key development over the standard version of DRT as presented in Section 6.2 is that whereas in DRT it is assumed that all discourse units are merged together in a single DRS, in SDRT each contribution to discourse is treated as a separate entity (a labelled DRSS), which has to be linked to the rest of the content tree via a rhetorical relation whose existence has to be inferred via defeasible inference.

For instance, consider the simple discourse in (6.14). Whereas in standard DRT both sentences would result in updates of the same DRS, in SDRT each sentence results in a new proposition. After the second sentence is semantically interpreted, resulting in DRS π_2 , discourse interpretation attempts to link

it with the existing content tree by finding a rhetorical relation linking it to discourse unit π_1 ; finding that π_2 may provide an Explanation for π_1 results in the interpretation in (6.15), whose basic tree structure is shown in (6.16).

(6.14) Max fell. John pushed him.



(6.16)

π_1
Explanation
 π_2

The interpretation of rhetorical relations is fully formalised in SDRT. Asher

and Lascarides introduce a distinction between three types of rhetorical relations. **Veridical** relations, which include Narration, Explanation, Elaboration, Background, Contrast and Parallel, are defined as the relations that satisfy a Satisfaction Schema requiring both propositions linked by the relation to be satisfied. Each of these relations then imposes some additional semantic constraints. For instance, $\text{Explanation}(\pi_1, \pi_2)$ entails $\text{cause}(e_{\pi_2}, e_{\pi_1})$. In addition, SDRT's repertoire of relations includes **Non-veridical** relations such as Alternation (SDRT's version of disjunction), in which only one of the relata is asserted to be true, and relations such as Correction, where the truth of the relata is mutually exclusive.

6.3.5 Connective-based theories of relational coherence

The problem identified by research on RST of identifying the set of rhetorical relations led Knott (1996) to propose a radical approach to the definition of such set: simply define it starting from the set of **cue phrases** signalling discourse relations—conjunctions like *while* in (6.17a), adverbials like *otherwise* in (6.17b), and prepositional phrases like (6.17c).¹⁹

- (6.17)a. Joshua eats Cheerios for breakfast, while Massimo eats muesli.
- b. Eat your Cheerios. Otherwise you're not going to watch Charlie and Lola.
- c. You've eaten your Cheerios every day this week. As a result, we can go to Go Bananas today.

Knott produced a list of over 200 such cue-phrases, and then proceeded to identify for each of them the discourse relations they expressed by means of substitution tests checking which cue phrases could replace other cue phrases in which context. In follow-up work, Webber *et al.* (2003) produced a drastically simplified version of Knott's taxonomy of relations by hypothesising a distinction between two types of cue phrases. According to Webber *et al.*, coordinating and subordinating conjunctions (*and*, *but*, *although*) express discourse relations;

but discourse adverbials such as *then*, *otherwise* or *instead* are anaphors. The resulting theory served as the basis for the annotation of the Penn Discourse Treebank (Webber *et al.*, 2005; Prasad *et al.*, 2008), at present the largest corpus annotated with information about relational coherence.

6.4 Salience

The anaphoric expressions discussed in Section 6.2 have different felicity conditions. Pronouns, for instance, can only be used felicitously to refer to discourse entities that ‘stand out,’ e.g., because they have been recently mentioned: e.g., Hobbs (1978a) reported that in his data, 90% of all pronoun antecedents were in the current sentence, and 98% in the current or the previous sentence, although there was no fixed distance beyond which no antecedent could be found (one pronominal antecedent was found 9 sentences back). This restriction of pronoun antecedents to the current and previous sentence for pronouns has been confirmed by every study of referential distance, if with slightly different figures: e.g., Hitzeman and Poesio (1998) found that around 8% of pronoun antecedents in their corpora were not in the current or previous sentence. But distance is less important for other types of anaphoric expressions: e.g., Givon (1992) found that 25% of definite antecedents in his study were in the current clause, 60% in the current or previous 20 clauses, but 40% were further apart. Vieira (1998) found that a window of 5 was optimal for definites.

These effects cannot simply be explained in terms of recency. For instance, there is a lot of evidence for a *first mention advantage*—a preference for pronouns and other anaphoric expressions to refer to entities mentioned in first position in the previous sentence (Gernsbacher and Hargreaves, 1988; Gordon *et al.*, 1993)—even though such entities are typically not the closest to the anaphoric expression. And researchers such as Grosz (1977), Linde (1979), Sanford and Garrod (1981), Gundel *et al.* (1993), and others, have argued that the production and interpretation of text in discourse is affected by attentional mechanisms

of the type found in visual interpretation: i.e., they claim that some parts of a discourse model (entities and / or propositions) are more **salient** than others, and that users of anaphoric expressions are sensitive to these differences. Consider for instance example (6.18) (from (Schuster, 1988)). As this example shows, personal pronouns seem to be used to refer to entities (in this case, the action of becoming a bum) that are 'more salient' than others; by contrast, demonstrative pronouns seem to be preferentially interpreted to refer to entities that, while salient, are not quite as salient.

(6.18)a. John thought about {becoming a bum}.

b. It would hurt his mother and it would make his father furious.

c. It would hurt his mother and that would make his father furious.

Similar differences between the felicity conditions of pronominal expressions are universal, and many discourse linguists have explained these effects by stipulating that discourses have one or more **topics**, and that languages in the world have special devices to refer to these topics: personal pronouns in English, *wa*-marked NPs in Japanese, and zero anaphors in Japanese and Romance languages such as Italian, Portuguese and Spanish (Givon, 1983; Gundel *et al.*, 1993; Grosz *et al.*, 1995).

Researchers such as Reichman (1985); Grosz and Sidner (1986); Albrecht and O'Brien (1993); Garrod (1994) made the further hypothesis that there are *two* types of salience effects: **local** effects such as those just discussed, which have to do with which entities are most salient at a given point in a conversation; and **global** effects, determined by the situational / topical or intentional structure of the text. We already discussed the right frontier constraint; the claim that discourses are segmented according to 'topics' or the episodic organization of the story is backed up by results such as those obtained by Anderson *et al.* (1983). Anderson and colleagues presented their subjects with a passage like in Figure 6.4, introducing a main character (in this case, female) and a secondary character (in this case, male) tied to the scenario. This first passage was followed

either by a sentence expressing immediate continuation of the episode (*Ten minutes later . . .*) or by one indicating that the story had moved on (*Ten hours later . . .*). Finally, the subjects were presented with either a sentence referring to the main entity, or to one referring to the scenario entity. Anderson *et al.* found an entity x delay effect: after the sentence expressing immediate continuation there was no difference in processing a pronoun referring to the main entity or a pronoun referring to the scenario entity, but when the text indicated a longer delay (and hence, a closure of the previous episode) the pronominal reference to the scenario entity was harder to process.

AT THE CINEMA

Jenny found the film rather boring.
 The projectionist had to keep changing reels.
 It was supposed to be a silent classic.
 a. Ten minutes later the film was forgotten
 Ten hours later the film was forgotten
 b. She was fast asleep
 c. He was fast asleep

Figure 6.4: The materials from (Anderson *et al.*, 1983)

In the rest of this Section we will discuss two highly influential theories of salience: the Activation Hierarchy theory proposed by (Gundel *et al.*, 1993), and the Grosz / Sidner theory of local and global focus in which some of the notions used by Gundel *et al.* are made (slightly) more precise.

6.4.1 Gundel et al's Activation Hierarchy

Gundel *et al.*'s theory of the conditions under which referring expressions are used (Gundel *et al.*, 1993) assumes that two factors interact in determining the choice of referring expression.

The first of these factors is the **activation hierarchy**: a speaker's choice of expression depends in part on assumptions about the 'cognitive status' of the referent in the hearer's information state. Gundel *et al.*'s 'activation levels' range from *type identifiability* for indefinite NPs, to *in focus* for pronouns.

in	> activated >	familiar >	uniquely >	referential >	type
focus			identifiable		identifiable
	<i>that</i>			indefinite	
<i>it</i>	<i>this</i>	<i>that N</i>	<i>the N</i>	<i>this N</i>	<i>a N</i>
	<i>this N</i>				

The second factor playing a role in Gundel *et al.*'s account are Grice's maxims of quantity:

Q1 Make your contribution as informative as possible

Q2 Do not make your contribution more informative than necessary

These maxims prevent the use of referring expressions associated with higher activation levels to refer to entities with a lower status, as we will see in a moment.

Thus for instance, according to Gundel *et al.* the reason for the contrast between the uses of personal pronouns and demonstrative pronouns highlighted by (6.18) is that using pronoun *that*²⁰ requires the referent to be **activated**, which status they characterise as "being represented in current short-term memory."²¹ This condition would also license the use of *this*-NPs to refer to entities in focus; what prevents this, according to Gundel *et al.*, is Grice's Q1: because a more specific referring form exists, the use of a demonstrative for entities in focus would implicate a lower activation level.

6.4.2 The Grosz / Sidner framework

The best-known theory of salience is the framework proposed by Grosz and Sidner (1986) and articulated in two levels: the **global focus** specifying the articulation of a discourse into segments, and the **local focus** of salience specifying how utterance by utterance the relative salience of entities changes.

Global Salience Grosz and Sidner's theory of the global focus is articulated around two main components: the **intentional structure** and the **focus spaces stack**. The intentional structure, discussed in Section 6.3.3, is an

intention-based account of the relational coherence of a discourse. Grosz and Sidner then propose that global salience has a stack structure: the (contents of) the intentions that dominate the intention associated with the present utterance are salient.

Other models of the global focus have also been proposed; in particular, Walker (1998) proposes a cache model for the global focus. The two models were compared by Poesio *et al.* (2006) in terms of the way they limit accessibility. Knott *et al.* (2001) argued that the intentional structure proposed by Grosz and Sidner, while perhaps appropriate for task-oriented dialogue, is not appropriate for many types of text.

Local Salience The second level of attention in Grosz and Sidner's theory is the so-called **local focus**. According to Grosz and Sidner and other researchers including Linde, Garrod and Sanford, and others, at every moment during a conversation or while reading text some entities are more salient than the others and are preferred antecedents for pronominalization and other types of anaphoric reference. Sidner (1979) proposed the first detailed theory of the local focus, articulated around two distinct foci: the **discourse focus**, meant to account for the phenomena normally explained in terms of the notion of 'discourse topic' (Gundel, 1974; Reinhart, 1981; Vallduvi, 1993) is usually introduced. In (6.19), the meeting with Ira is the discourse focus and serves as privileged antecedent for certain types of anaphoric reference.

(6.19)a. I want to schedule a meeting with Ira.

b. It should be at 3p.m.

c. We can get together in his office

Sidner also introduces an **actor focus**, supposed to capture some of the effects accounted for in previous theories through subject assignment, such (6.20).

(6.20) John gave a lot of work to Bill. He often helps friends this way.

According to Sidner, the local focus changes after every sentence as a result of mention and coreference. Extremely complex algorithms are provided for both foci and for their use for anaphoric reference.

Centering theory (Grosz *et al.*, 1995) was originally proposed as just a simplified version of Sidner’s theory of the local focus (Grosz *et al.*, 1983) but eventually it evolved in a theory of its own –in fact, the dominant paradigm for theorizing about salience in computational linguistics and, to some extent, in psycholinguistics and corpus linguistics as well (see, e.g., the papers in Walker *et al.* (1998)). According to Centering, every **utterance** updates the local focus by introducing new **forward looking centers** (mentions of discourse entities) and updating the focal structure. Forward looking centers are **ranked**: this means that each utterance has a most highly ranked entity, called **Preferred Center** (CP), which corresponds broadly to Sidner’s actor focus. In addition, Centering hypothesises the existence of an object playing the role of the discourse topic or discourse focus: the **backward looking center**, defined as follows:

Constraint 3 $CB(U_i)$, the **Backward-Looking Center** of utterance U_i , is the highest ranked element of $CF(U_{i-1})$ that is realised in U_i .

Several psychological experiments have been dedicated to testing the claims of Centering, and in particular those concerning pronominalization, known as Rule 1:

Rule 1 If any CF in an utterance is pronominalised, the CB is.

Hudson-D’Zmura and Tanenhaus (1998) found a clear preference for subjects, which could however also be accounted for in terms of subject assignment. Gordon and colleagues carried out a series of experiments that, they argued, demonstrated certain features of the theory. Gordon *et al.* (1993), for instance, revealed a *repeated name penalty*—a preference for avoiding repeating full names when an entity is mentioned in subject or first mention position, and using pronouns instead. Thus for instance Gordon *et al.* found an increase in reading

time when processing sentences b–c of (6.21), with respect to reading sentences b–c of ex:RNP:2 in which the proper name in subject position *Bruno* has been replaced by pronoun *He*.

(6.21)a. Bruno was the bully of the neighborhood.

- b. Bruno chased Tommy all the way home from school one day.
- c. Bruno watched Tommy hide behind a big tree and start to cry.
- d. Bruno yelled at Tommy so loudly that the neighbors came outside.

(6.22)a. Bruno was the bully of the neighborhood.

- b. He chased Tommy all the way home from school one day.
- c. He watched Tommy hide behind a big tree and start to cry.
- d. He yelled at Tommy so loudly that the neighbors came outside.

Poesio *et al.* (2004) carried out a systematic corpus-based investigation of the claims of Centering, that revealed among other things that entity coherence between utterances is much less strong than expected, so that the majority of utterances do not have a CB.

The main alternative to Grosz and Sidner's **discrete** models of salience are **activation-based** models in which there is no fixed number of foci, but in which all entities have a level of activation (Klapholz and Lockman, 1975; Alshawi, 1987; Lappin and Leass, 1994; Strube, 1998; Tetreault, 2001).

Further readings and relevant resources

Regarding resources, we mentioned available corpora for studying the aspects of discourse we discussed in connection with the specific topics.

As far as readings on anaphora and entity coherence are concerned, the most extensive treatments currently available are (Kamp and Reyle, 1993) for the linguistic of anaphora and discourse models, (Garnham, 2001) for the psychological evidence, and (Mitkov, 2002) for the computational side, but all three books are beginning to be quite outdated. A new textbook we wrote on the topic (Poe-

sio *et al.*, pear) is in press and will come out by December 2015. The papers (Knott *et al.*, 2001; Poesio *et al.*, 2004) are also recommended; the latter paper also covers local salience.

An excellent textbook by Manfred Stede covers a lot of ground on discourse structure and relational coherence (Stede, 2011). On discourse structure and relational coherence, the most highly recommended text for the linguistic background is still Fox's *Discourse Structure and Anaphora* (Fox, 1987). For discourse processing from a psychological perspective, we recommend (Graesser *et al.*, 1997). The survey by Taboada and Mann (2006) covers not just RST but more in general many issues with discourse structure and relational coherence and is highly recommended. The survey by Webber *et al.* (2011) also covers a number of approaches to discourse parsing.

On salience, we recommend three papers: (Grosz and Sidner, 1986) on global and local salience, (Grosz *et al.*, 1995) on Centering, and (Gundel *et al.*, 1993) for a linguistic perspective.

Bibliography

- Albrecht, Jason E. and O'Brien, Edward J (1993). Updating a mental model: maintaining both local and global coherence. *Journal Experimental Psychology: Learning Memory and Cognition*, **19**, 1061–1070.
- Alshawi, Hiyan (1987). *Memory and Context for Language Interpretation*. Cambridge University Press, Cambridge.
- Anderson, Anne, Garrod, Simon C., and Sanford, Anthony (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology*, **35**, 427–440.
- Asher, Nicholas (1993). *Reference to Abstract Objects in English*. D. Reidel, Dordrecht.
- Asher, Nicholas and Lascarides, Alex (2003). *The Logic of Conversation*. Cambridge University Press.
- Barwise, Jon and Perry, John (1983). *Situations and Attitudes*. The MIT Press.
- Carlson, Lynn, Marcu, Daniel, and Okurowski, Mary E. (2003). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In J. Kuppevelt and R. Smith, editors, *Current Directions in Discourse and Dialogue*, pages 85–112. Kluwer.
- Clark, Herbert H. (1977). Bridging. In P. N. Johnson-Laird and P. Wason, editors, *Thinking: Readings in Cognitive Science*, pages 411–420. Cambridge University Press, London and New York.

- Dowty, David R. (1986). The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy*, **9**(1).
- Fox, Barbara A. (1987). *Discourse Structure and Anaphora*. Cambridge University Press, Cambridge, UK.
- Garnham, Allan (1982). On-line construction of representations of the content of texts. Reproduced by Indiana University Linguistics Club.
- Garnham, Allan (2001). *Mental models and the interpretation of anaphora*. Psychology Press.
- Garrod, Simon C. (1994). Resolving pronouns and other anaphoric devices: The case for diversity in discourse processing. In C. Clifton, L. Frazier, and K. Rayner, editors, *Perspectives in Sentence Processing*. Lawrence Erlbaum.
- Gernsbacher, Morton A. and Givon, T., editors (1995). *Coherence in Spontaneous Text*. John Benjamins.
- Gernsbacher, M. A. and Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, **27**, 699–717.
- Givon, T., editor (1983). *Topic continuity in discourse : a quantitative cross-language study*. J. Benjamins, Amsterdam and Philadelphia.
- Givon, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, **30**.
- Gordon, Peter C., Grosz, Barbara J., and Gillion, Laura A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, **17**, 311–348.
- Graesser, Arthur C., Millis, Keith K., and Zwaan, Rolf A. (1997). Discourse comprehension. *Annual Review Psychology*, **48**, 163–189.

- Grice, H. Paul (1975). Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and Semantics, vol.3: Speech Acts*, pages 41–58. Academic Press, New York.
- Grimes, Joseph E. (1972). *The Thread of Discourse*. Cornell University Press, Ithaca.
- Groenendijk, Jeroen and Stokhof, Martin (1991). Dynamic Predicate Logic. *Linguistics and Philosophy*, **14**, 39–100.
- Grosz, Barbara, Joshi, Araving, and Weinstein, Scott (1983). Providing a unified account of definite noun phrases in discourse. In *Proc. ACL-83*, pages 44–50, Cambridge, MA.
- Grosz, Barbara J. (1977). *The Representation and Use of Focus in Dialogue Understanding*. Ph.D. thesis, Stanford University.
- Grosz, Barbara J. and Sidner, Candace L. (1986). Attention, intention, and the structure of discourse. *Computational Linguistics*, **12**(3), 175–204.
- Grosz, Barbara J., Joshi, Aravind K., and Weinstein, Scott (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, **21**(2), 202–225. (The paper originally appeared as an unpublished manuscript in 1986.).
- Gundel, Jeanette K. (1974). *The Role of Topic and Comment in Linguistic Theory*. Ph.D. thesis, University of Texas at Austin. Reprinted by Garland Publishing, New York and London, 1988.
- Gundel, Jeanette K., Hedberg, Nancy, and Zacharski, Ron (1993). Cognitive status and the form of referring expressions in discourse. *Language*, **69**(2), 274–307.
- Halliday, Michael A. K. and Hasan, Ruqaiya (1976). *Cohesion in English*. Longman, London.

- Heim, Irene (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts at Amherst.
- Heim, Irene (1983). File change semantics and the familiarity theory of definiteness. In R. Bauerle, C. Schwarze, and A. von Stechow, editors, *Meaning, Use and Interpretation of Language*. de Gruyter, Berlin.
- Hengeveld, Kees and Mackenzie, J. Lachlan (2008). *Functional Discourse Grammar*. Oxford.
- Hitzeman, Janet and Poesio, Massimo (1998). Long-distance pronominalisation and global focus. In *Proc. of ACL/COLING, vol. 1*, pages 550–556, Montreal.
- Hobbs, Jerry R. (1978a). Resolving pronoun references. *Lingua*, **44**, 311–338.
- Hobbs, Jerry R. (1978b). Why is discourse coherent? Technical Note 176, SRI International.
- Hobbs, Jerry R. (1979). Coherence and coreference. *Cognitive Science*, **3**, 67–90.
- Hobbs, Jerry R. (1985). On the coherence and structure of discourse. Report CSLI-85-37, CSLI.
- Hobbs, Jerry R., Stickel, Mark, Appelt, David, and Martin, Paul (1993). Interpretation as abduction. *Artificial Intelligence Journal*, **63**, 69–142.
- Hovy, Eduard H. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence*, **63**(1-2), 341–385.
- Hovy, Eduard H. and Maier, Elisabet (1991). Parsimonious and profligate approaches to the question of discourse structure relations. Unpublished manuscript, available from <http://www.isi.edu/natural-language/people/hovy/papers/93discproc.pdf>.
- Hudson-D’Zmura, Susan and Tanenhaus, Michael K. (1998). Assigning antecedents to ambiguous pronouns: The role of the center of attention as the

- default assignment. In M. A. Walker, A. K. Joshi, and E. F. Prince, editors, *Centering in Discourse*, pages 199–226. Oxford University Press.
- Kamp, Hans (1979). Events, instant and temporal reference. In R. Bauerle, U. Egli, and A. von Stechow, editors, *Semantics from Different Points of View*, pages 376–417. Springer-Verlag.
- Kamp, Hans (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*. Mathematical Centre, Amsterdam.
- Kamp, Hans and Reyle, Uwe (1993). *From Discourse to Logic*. D. Reidel, Dordrecht.
- Karttunen, Lauri (1976). Discourse referents. In J. McCawley, editor, *Syntax and Semantics 7 - Notes from the Linguistic Underground*, pages 363–385. Academic Press, New York.
- Kehler, Andrew (2002). *Coherence, Reference, and the Theory of Grammar*. CSLI.
- Kehler, Andrew, Kertz, Laura, Rohde, Hannah, and Elman, Jeffrey (2008). Coherence and coreference revisited. *Journal of Semantics*, **25**(1), 1–44.
- Kibble, Rodger (2001). A reformulation of Rule 2 of Centering Theory. *Computational Linguistics*, **27**(4), 579–587.
- Kintsch, Walter and van Dijk, Teun (1978). Towards a model of discourse comprehension and production. *Psychological Review*, **85**, 363–394.
- Klapholz, David and Lockman, Abraham (1975). Contextual reference resolution. *American Journal of Computational Linguistics*.
- Knott, Alistair (1996). *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Ph.D. thesis, University of Edinburgh, Department of Artificial Intelligence.

- Knott, Alistair, Oberlander, John, O'Donnell, Mick, and Mellish, Chris (2001). Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord, and W. Spooren, editors, *Text representation: linguistic and psycholinguistic aspects*, pages 181–196. John Benjamins, Amsterdam and Philadelphia.
- Lappin, Shalom and Leass, Herbert J. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, **20**(4), 535–562.
- Lascarides, Alex and Asher, Nicholas (2007). Segmented discourse representation theory: Dynamic semantics with discourse structure. In H. Bunt and R. Muskens, editors, *Computing Meaning*, volume 3, pages 87–124. Springer.
- Lesgold, Alan, Lajoie, Susanne, Bunzo, Marilyn, and Eggan, Gary (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin and R. Chabay, editors, *Computer assisted instruction and intelligent tutoring systems: Shared issues and complementary approaches*, pages 201–238. Erlbaum, Hillsdale, NJ.
- Linde, Charlotte (1979). Focus of attention and the choice of pronouns in discourse. In T. Givon, editor, *Syntax and Semantics 12*. Academic Press.
- Mann, William C. (1984). Discourse structures for text generation. Technical Report ISI/RR-84-127, USC, Information Sciences Institute.
- Mann, William C. and Thompson, Sandra A. (1983). Relational propositions in discourse. Technical Report ISI/RR-83-115, USC, Information Sciences Institute.
- Mann, William C. and Thompson, Sandra A. (1988). Rhetorical Structure Theory: Towards a functional theory of text organization. *Text*, **8**(3), 243–281.
- Mann, William C. and Thompson, Sandra A., editors (1992). *Discourse de-*

- scription : diverse linguistic analyses of a fund-raising text.* J. Benjamins, Amsterdam, Philadelphia.
- Marcu, Daniel (2000). *The Theory and Practice of Discourse Parsing and Summarization.* MIT Press, Cambridge, MA.
- Marcu, Daniel, Romera, Magdalena, and Amorrortu, Estibaliz (1999). Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In *Workshop on Levels of Representation in Discourse*, pages 71–78. University of Edinburgh.
- Meyer, Bonnie (1975). *The Organization of Prose and its Effects on Memory.* North-Holland, Amsterdam.
- Mitkov, Ruslan (2002). *Anaphora Resolution.* Longman.
- Moore, Johanna and Pollack, Martha (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics*, **18**(4), 537–544.
- Moser, Megan and Moore, Johanna D. (1996a). On the correlation of cues with discourse structure: Results from a corpus study. Unpublished manuscript.
- Moser, Megan and Moore, Johanna D. (1996b). Toward a synthesis of two accounts of discourse structure. *Computational Linguistics*, **22**(3), 409–419.
- Muskens, Reinhard A. (1996). Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy*, **19**, 143–186.
- O’Donnell, Mick, Mellish, Chris, Oberlander, Jon, and Knott, Alistair (2001). ILEX: An architecture for a dynamic hypertext generation system. *Natural Language Engineering*, **7**, 225–250.
- Partee, Barbara H. (1973). Some structural analogies between tenses and pronouns in English. *Journal of Philosophy*, **70**, 601–609.
- Poesio, Massimo (2004). Discourse annotation and semantic annotation in the GNOME corpus. In *Proc. of the ACL Workshop on Discourse Annotation*, pages 72–79, Barcelona.

- Poesio, Massimo and Vieira, Renata (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, **24**(2), 183–216. Also available as Research Paper CCS-RP-71, Centre for Cognitive Science, University of Edinburgh.
- Poesio, Massimo, Stevenson, Rosemary, Di Eugenio, Barbara, and Hitzeman, Janet M. (2004). Centering: A parametric theory and its instantiations. *Computational Linguistics*, **30**(3), 309–363.
- Poesio, Massimo, Patel, Amrita, and Di Eugenio, Barbara (2006). Discourse structure and anaphora in tutorial dialogues: an empirical analysis of two theories of the global focus. *Research in Language and Computation*, **4**, 229–257. Special Issue on Generation and Dialogue.
- Poesio, Massimo, Stuckardt, Roland, and Versley, Yannick (To Appear). *Anaphora Resolution: Algorithms, Resources and Applications*. Springer, Berlin. To appear.
- Polanyi, Livia (1985). A theory of discourse structure and discourse coherence. In *Proc. Chicago Linguistic Society*, pages 306–322.
- Polanyi, Livia (1995). The linguistic structure of discourse. Unpublished manuscript, available online as http://homepage.mac.com/livia.polanyi/publications/lm_CSLI_1995.pdf.
- Prasad, Rashmi, Dinesh, Nikhal, Lee, Alan, Miltsakaki, Eleni, Robaldo, Livio, Joshi, Aravind, and Webber, Bonnie (2008). The penn discourse treebank 2.0. In *Proc. of LREC*.
- Reichman, Rachel (1985). *Getting Computers to Talk Like You and Me*. The MIT Press, Cambridge, MA.
- Reinhart, Tanya (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, **27**(1). Also distributed by Indiana University Linguistics Club.

- Rooth, Mats (1987). Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics. In P. Gärdenfors, editor, *Generalized Quantifiers*, pages 237–268. D. Reidel, Dordrecht, The Netherlands.
- Rosé, Carolyn P., Di Eugenio, Barbara, Levin, Lori S., and Van Ess-Dykema, Carol (1995). Discourse processing of dialogues with multiple threads. In *Proc. ACL*, Boston, MIT.
- Sanders, Ted and Spooren, Wilbert (2001). Text representation as an interface between language and its users. In T. Sanders, J. Schilperoord, and W. Spooren, editors, *Text Representation: Linguistic and Psycholinguistic Aspects*. John Benjamins.
- Sanders, T., Spooren, Wilbert, and Noordman, Leo G. M. (1992). Towards a taxonomy of coherence relations. *Discourse Processes*, **15**, 1–35.
- Sanford, Anthony J. and Garrod, Simon C. (1981). *Understanding Written Language*. Wiley, Chichester.
- Scha, Remko and Polanyi, Livia (1988). An augmented context free grammar for discourse. In *Proceedings of the 12th Conference on Computational Linguistics*, pages 573–577, Prague.
- Schuster, Ellen (1988). Pronominal reference to events and actions: Evidence from naturally-occurring data. LINC LAB 100, University of Pennsylvania, Dept. of Computer and Information Science, Philadelphia.
- Sidner, Candace L. (1979). *Towards a computational theory of definite anaphora comprehension in English discourse*. Ph.D. thesis, MIT.
- Sperber, Dan and Wilson, Deirdre (1986). *Relevance: Communication and Cognition*. Harvard University Press, Cambridge, MA.
- Stede, Manfred (2011). *Discourse Processing*. Synthesis Lectures on Human Language Technology. Morgan&Claypool.

- Strube, Michael (1998). Never look back: An alternative to centering. In *Proc. of COLING-ACL*, pages 1251–1257, Montreal.
- Taboada, Maite and Hadic Zabala, L. (2008). Deciding on units of analysis within centering theory. *Corpus Linguistics and Linguistic Theory*, **4**(1), 63–108.
- Taboada, Maite and Mann, William C. (2006). Rhetorical structure theory: looking back and moving ahead. *Discourse Studies*, **8**(3), 423–459.
- Tetreault, Joel R. (2001). A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*, **27**(4), 507–520.
- Vallduvi, Enric (1993). Information packaging: a survey. Research Paper RP-44, University of Edinburgh, HCRC.
- van Deemter, Kees and Kibble, Rodger (2000). On coreferring: Coreference in MUC and related annotation schemes. *Computational Linguistics*, **26**(4), 629–637. Squib.
- van Dijk, Teun A. (1977). *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. Longman.
- van Dijk, Teun A. and Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.
- Vieira, Renata (1998). *Definite Description Resolution in Unrestricted Texts*. Ph.D. thesis, University of Edinburgh, Centre for Cognitive Science.
- Walker, Marilyn A. (1998). Centering, anaphora resolution, and discourse structure. In M. A. Walker, A. K. Joshi, and E. F. Prince, editors, *Centering in Discourse*, chapter 19, pages 401–435. Oxford University Press.
- Walker, Marilyn A., Joshi, Aravin K., and Prince, Ellen F., editors (1998). *Centering Theory in Discourse*. Clarendon Press / Oxford.

- Webber, Bonnie, Joshi, Aravind, Miltsakaki, Eleni, Prasad, Rashmi, Dinesh, Nikhil, Lee, Alan, and Forbes, Katherine (2005). A short introduction to the penn discourse treebank. *Copenhagen Working Papers in Linguistics*.
- Webber, Bonnie, Egg, Markus, and Kordoni, Valia (2011). Discourse structure and language technology. *Natural Language Engineering*.
- Webber, Bonnie L. (1979). *A Formal Approach to Discourse Anaphora*. Garland, New York.
- Webber, Bonnie L. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, **6**(2), 107–135.
- Webber, Bonnie L., Stone, Matthew, Joshi, Aravind, and Knott, Alistair (2003). Anaphora and discourse structure. *Computational Linguistics*, **29**(4), 545–587.
- Wolf, Florian and Gibson, Edward (2006). *Coherence in Natural Language: Data Structures and Applications*. The MIT Press.
- Zwaan, Rolf A., Langston, Mark C., and Graesser, Arthur C. (1995). The construction of situation models in narrative comprehension: an event-indexing model. *Psychological Science*, **6**(292–197).

Notes

¹Indeed many linguists proposed ‘discourse grammars’ very similar to sentence grammars, i.e., with fixed rules (van Dijk, 1977; Scha and Polanyi, 1988), although it is now generally accepted that the factors governing coherence are pragmatic in nature (Grice, 1975; Sperber and Wilson, 1986) so even the linguists still using the term ‘grammar’ tend to use the term ‘rule’ to refer to preferences (Hengeveld and Mackenzie, 2008).

²For more on anaphora, see Chapter 27.

³Some cases of reference are best viewed as depending on what is usually called the **discourse situation** or **utterance situation** (Barwise and Perry, 1983), that includes both the linguistic context and the surroundings in which the participants operate.

⁴An extreme abstraction!

⁵<http://projects ldc.upenn.edu/ace/data>

⁶<http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T24>

⁷<http://www.anaphoricbank.org>

⁸<http://clic.ub.edu/ancora/>

⁹<http://ufal.mff.cuni.cz/pdt2.0/>

¹⁰<http://www.clips.ua.ac.be/~iris/corea.html>

¹¹<http://www.sfs.uni-tuebingen.de/tuebadz.shtml>

¹²<http://www.anaphoricbank.org>

¹³<http://cl.naist.jp/nldata/corpus/>

¹⁴The distinction between intentional and informational coherence is implicitly present in most theories of relational coherence but is made more explicit in Relational Discourse Analysis (Moser and Moore, 1996b).

¹⁵For a particularly insightful discussion of discourse units, see (Polanyi, 1995). See also (Poesio *et al.*, 2004; Taboada and Hadic Zabala, 2008).

¹⁶E.g., Taboada and Mann (2006) state that “... the (RST) analyst seeks to find an annotation that includes every part of the text in one connected whole” (p. 425).

¹⁷<http://www.seas.upenn.edu/~pdtb/>

¹⁸The analysis can be found on the RST official website, <http://www.sfu.ca/rst/>, which contains several RST analyses of texts of various length.

¹⁹These are slightly modified versions of examples from (Webber *et al.*, 2005).

²⁰But not full *that* NPs, which only require the referent to have the lower ‘familiar’ status.

²¹In fact, for demonstrative nominals, Gundel *et al.* claim that the referent has to be speaker-active—introduced by the speaker.