ELSEVIER

Journal of Pragmatics 29 (1998) 57–74

journal of PRAGMATICS

Discussion

Discourse, coherence and relevance: A reply to Rachel Giora[☆]

Deirdre Wilson

Department of Linguistics, University College London, Gower Street, London WC1E 6BT, UK

Received December 1996

Abstract

This paper is a response to the arguments of Rachel Giora (1997) that relevance theory, as developed by Sperber and Wilson (1986/1995), cannot account for intuitions of text or discourse coherence. The paper has three main aims: first, to clarify the goals and mechanisms of relevance theory; second, to show how relevance theory can shed some light on intuitions of local and global coherence; and third, to compare the resulting relevance-theoretic account with Giora's coherence-based account.

1. Introduction

Rachel Giora (1997) raises some important questions about the interpretation and evaluation of texts. Her main point is a critical one: relevance theory, as developed by Sperber and Wilson (1986/1995), cannot account for intuitions of text or discourse coherence. She also sketches a more positive account, based on notions of topic-relevance and informativeness and developed in more detail in Giora (1985a,b, 1988). I would like to respond briefly to her critical remarks, and go on to consider the more general question of how the interpretation and evaluation of texts is best approached.

2. The problem

Let me start with a question. What is a theory of text or discourse coherence designed to do? Many coherence theorists set themselves two related goals: (a) to

☆ I would like to thank Dan Sperber and the members of the University College London Pragmatics Reading Group for comments on an earlier version.

PII S0378-2166(97)00012-X

provide a theory of comprehension, explaining how discourses are understood; (b) to provide a theory of evaluation, explaining intuitions of discourse well-formedness, acceptability or appropriateness.[1] Goal (a) is shared with relevance theory, which aims to provide a theory of comprehension. However, relevance theory has no explicit goal equivalent to (b).

Giora's interest is in goal (b). Her aim is to account for intuitions of coherence and degrees of coherence, which she describes at various points in her paper as intuitions about the 'well-formedness', 'appropriateness', 'acceptability' or 'desirability' of texts. These notions are obviously not equivalent: well-formedness exists only relative to a set of well-formedness rules; appropriateness, acceptability and desirability are relative to individuals, situations and contexts. What they all have in common is that they relate to the evaluation of texts. Giora makes no commitment to developing a coherence-based theory of comprehension: apart from noting that coherent texts are the easiest to process,[2] her concern is with evaluation throughout.

The main aim of relevance theory in the domain of verbal communication is to explain how utterances are understood. As a by-product, it claims to shed light on certain intuitions of acceptability or unacceptability, for which coherence theorists have sometimes offered alternative accounts. However, there is no commitment to the existence of well-defined types of text or discourse, each with its own well-formedness conditions. Indeed, most relevance theorists would see it as an advantage of the theory that it allows for a more flexible, open-ended view of discourse than talk of well-formedness conditions suggests.

Giora's objection to relevance theory is that it does not provide an adequate account of discourse evaluation. In particular, she suggests that it sheds no light on two main types of intuition: (a) local intuitions about when two adjacent discourse segments are related (to each other or to a common topic); (b) global intuitions about when a text or discourse 'hangs together' as a unit, rather than constituting a series of isolated remarks. This is the substantive issue I want to address. First, though, I need to clear up some misunderstandings about how relevance theory works.

3. Misconceptions

3.1. Maximal relevance, optimal relevance and the criterion of consistency with the principle of relevance

Giora's main objection to relevance theory is that it cannot account for intuitions of discourse coherence, because:

(1) a. A discourse may be 'S&W relevant' [i.e. relevant in Sperber and Wilson's sense] to an individual, but nonetheless judged incoherent;

---

[1] See, for example, Hobbs (1978, 1990); Hobbs et al. (forthcoming); Mann and Thompson (1986); Samet and Schank (1984).

[2] On the role of coherence in facilitating processing, see also Giora (1996).

b. A discourse may be 'S&W irrelevant' to an individual, but nonetheless judged coherent.

In other words, 'S&W relevance' is neither necessary nor sufficient for coherence.

To assess this objection, we need to know what Giora means by *S and W relevance*.[3] In *Relevance* (Sperber and Wilson, 1986/1995) Dan Sperber and I define several notions of relevance: *relevance in a context*, *relevance to an individual*, *relevance of a phenomenon*, *maximal relevance* and *optimal relevance*. We propose two relevance-based principles, one governing cognition and the other communication:

(2) *Cognitive principle of relevance*
Human cognition tends to be geared to the maximisation of relevance.

(3) *Communicative principle of relevance*
Every act of ostensive communication communicates a presumption of its own optimal relevance.

In the first edition of *Relevance*, although both principles were present, only the second was named; the result was that many readers failed to notice that two separate principles were involved. What we called 'the principle of relevance' was the communicative principle of optimal relevance; however, many commentators confused this with the cognitive principle of maximal relevance. Giora is one of these. Although her concern is with communication rather than cognition, throughout her article, when she talks of 'S&W relevance', she means *maximal relevance*, and when she talks of the 'principle of relevance', she means the cognitive principle in (2) above, rather than the communicative principle in (3).[4]

In *Relevance* and elsewhere, Dan Sperber and I specifically deny that *communication* is geared to the maximisation of relevance. Communicators are not always expected to give the most relevant possible information, or to present it in the least effort-demanding way, as a principle of maximal relevance would suggest. On the content side, the speaker may not have the information that the hearer would find most relevant; she may be unwilling to give it, or unable to think of it at the time. On the formal side, lack of time, lack of ability or stylistic preferences may prevent her expressing herself in the most economical way. Moreover, the most relevant utterance the speaker can think of may still not be relevant enough to be worth the hearer's attention; we need to set a lower limit on the expected degree of relevance, to explain why such an utterance would generally not be produced.

---

[3] Giora oddly describes Sperber and Wilson's notion of relevance as 'more recent' than those of Reinhart (1981); Brown and Yule (1983); Giora (1985a,b); Berg (1991). The earliest publications in relevance theory were in 1979, and there is a very substantial literature predating the publication of *Relevance* in 1986.

[4] In the Postface to the Second Edition of *Relevance* (Sperber and Wilson, 1986/1995) we distinguish explicitly between a First, or Cognitive Principle of Relevance and a Second, or Communicative Principle of Relevance, as in (2) and (3). For discussion of other common misunderstandings of relevance theory, see Sperber and Wilson (1987, section 3).

The communicative principle of relevance in (3), and the notion of optimal relevance in (4), are designed to take these factors into account:

(4) *Optimal relevance*
An ostensive stimulus (e.g. an utterance), on a given interpretation, is optimally relevant iff:
a. It is relevant enough for it to be worth the addressee's effort to process it;
b. It is the most relevant one compatible with the communicator's abilities and preferences.

Clause (a) sets a lower limit on the expected degree of relevance. Clause (b) takes the speaker's abilities and preferences into account. More positively, clause (b) incorporates the idea that if the speaker can see a way of increasing the relevance of the utterance at no cost to herself, it is in her interest to do so: the more relevant the utterance, the more likely the hearer will be to attend to it and understand it correctly.

Giora's objection in (1) should therefore be restated, substituting the communicative notion of optimal relevance for the cognitive notion of maximal relevance:

(5) a. A discourse may be optimally relevant to an individual, but nonetheless judged incoherent;
b. A discourse which is not optimally relevant to an individual may nonetheless be judged coherent.

In fact, there is nothing in (5) that Sperber and Wilson would disagree with. Although the notion of optimal relevance is central to our account of communication, we nowhere claim that an utterance is acceptable on a given interpretation if and only if it is optimally relevant on that interpretation. Giora's paper does raise a genuine issue, but to see what it is, we need to restate the objection once more.

Relevance theory aims to explain how the addressee decides on the intended interpretation of an utterance. According to the communicative principle of relevance, every utterance (or other act of ostensive communication) creates an expectation of optimal relevance. However, it does not follow that this expectation is always satisfied, in a way the speaker could manifestly have foreseen. In the first place, there are cases of accidental irrelevance, where an utterance is irrelevant for reasons the speaker manifestly could not have foreseen. In *Relevance* (pp. 159–160) we consider the following example. Mary knows that Peter buys all Iris Murdoch's books. She sees the latest one being put on display in the bookshop, and rushes to tell him:

(6) Iris Murdoch's new book is in the bookshops.

In fact, Peter has already seen the book, and the information in (6) is irrelevant to him. This would not make Mary's utterance any less appropriate or acceptable. More generally, an utterance will be acceptable on a given interpretation as long as the hearer can see how the speaker might reasonably have *expected* it to be optimally relevant to him.

In the second place, there are cases of accidental relevance, where an utterance is optimally relevant on a given interpretation, but the speaker manifestly could not have expected it to be relevant in that way. Here is an example. Susan, who plays for the school cricket team, has a collection of unusual pets (slow-worms, stick insects, bats). Sometimes she brings them to school to tease the games teacher, who is known to have a phobia for bats. One day, the teacher says to Susan and her classmates:

(7) Please leave your bats at home tomorrow.

As it happens, the first interpretation of (7) to occur to Susan is that the teacher is asking her to leave her pet bats at home. The utterance on this interpretation would be optimally relevant to Susan; indeed, if the teacher had been speaking to her alone, it is one he might reasonably have produced. However, since he was clearly addressing the group of classmates as a whole, he could not reasonably have expected his utterance, on this interpretation, to be optimally relevant to them. Susan should therefore reject this interpretation in favour of the one on which the teacher is telling the class to leave their cricket bats at home.

To allow for these two types of case, we claim in *Relevance* that interpretations are accepted or rejected according to the following criterion of *consistency with the principle of relevance*:[5]

(8) *Consistency with the (communicative) principle of relevance*
An utterance, on a given interpretation, is consistent with the (communicative) principle of relevance if and only if the speaker might reasonably have expected it to be optimally relevant on that interpretation.

Giora's objection should therefore be restated as follows:

(9) a. A discourse (on a given interpretation) may be judged consistent with the (communicative) principle of relevance, but nonetheless incoherent;
b. A discourse (on a given interpretation) may be judged inconsistent with the (communicative) principle of relevance, but nonetheless coherent.

This raises a genuine issue, and I will return to it in section 4.

3.2. Relevance in a context and relevance to an individual

In *Relevance*, we introduce our technical concept of relevance in several stages. First, we define the notion of *relevance in a context*, for a given assumption and a given context; the greater the contextual effects achieved by processing that assumption in that context, and the smaller the processing effort required, the greater the relevance (p. 125). Next, we define a notion of *relevance to an individual*, for a given

[5] For further discussion of these issues, see Sperber (1994).

assumption and an individual with access to a variety of contexts. Possible choices of context differ in the processing effort required and the contextual effects achieved; we claim that the choice of context is determined by the search for maximal relevance (p. 145). Finally, we define a notion of *relevance of a phenomenon to an individual*. A phenomenon makes manifest a variety of assumptions; an individual has access to a variety of contexts. The aim of the individual is to choose the best possible combination of assumption and context; we claim that the choice is again governed by the search for maximal relevance (pp. 152–153).

At stage 1, illustrating the notion of *relevance in a context*, we showed that given the context in (10a–c), the assumption in (11a) would be more relevant than those in (11b) and (11c):

(10) a. People who are getting married should consult a doctor about possible hereditary risks to their children.
b. Two people both of whom have thalassemia should be warned against having children.
c. Susan has thalassemia.

(11) a. Bill, who has thalassemia, is getting married to Susan.
b. Susan, who has thalassemia, is getting married to Bill.
c. Bill, who has thalassemia, is getting married to Susan, and 1967 was a great year for French wines.

(11a) is more relevant than (11b) on the contextual-effect side: it has a contextual implication which (11b) lacks. (11a) is more relevant than (11c) on the processing-effort side: (11c) contains additional material which is more costly to process, but yields no extra effects. We stressed the artificiality of these examples, which were designed purely to illustrate the technical notion of relevance in a context, and added:

> "Readers should try to resist the tendency to supply much richer and more appropriate contexts, a tendency which will be discussed at length later on." (Sperber and Wilson, 1986: 125)

Giora seems to have been unable to resist the temptation. She objects to our analysis of (11a–c) on the ground that:

> "Were contexts indeed given, this would be an adequate analysis. However, since contexts, we are told, are not given, but are rather chosen and determined by the search for S&W relevance, an alternative analysis is required. Such analysis will show that both [(11a)] and [(11b)] are equally S&W relevant." (Giora, 1997: 20)

Yet at the stage where these examples were introduced, we were not making an empirical claim about the relevance of (11a–c) to an individual with a choice of contexts, but merely showing how the relevance of a given assumption in a given context depends on the twin factors of contextual effects and processing effort. The notion of relevance to an individual, and the idea that contexts are constructed in the search for relevance, was discussed at length in *Relevance* chapter 3, sections 3 and

4. As for the communicative notion of optimal relevance, which would be the one to use in assessing the acceptability of *utterances* of (11a–c), it had not even been introduced at the point where we were discussing these examples, which were described as *assumptions*, not utterances, throughout.

3.3. Relevance and importance

Having rejected our argument that the choice of context is determined by the search for relevance, Giora considers an alternative account, in which the notions of 'importance' or 'interest' play a central role:[6]

> "[W]hat guides the choice of the appropriate context is what interlocutors consider most 'important' to them – what 'matters' to them, 'interests' or 'concerns' them." (Giora, 1997: 21–22)

She goes on to reject the more general relevance-theoretic claim that attention goes to information that seems relevant, and suggests that it goes to information that seems 'interesting' or 'important' instead. These suggestions involve the idea that relevance and importance or interest can come apart in such a way that one phenomenon – say, an utterance – may be simultaneously less relevant (in our sense) and more interesting than another, and vice versa. But this idea seems unfounded.[7]

Giora uses the following example to show that relevance and importance or interest can come apart. Someone walks into a lecture and announces:

(12) Ladies and gentlemen, I have to tell you that the building's on fire.

This example was used in Smith and Wilson (1992: 5) to show how the notion of optimal relevance can help with reference assignment. We claimed that if 'the building' in (12) is taken to refer to the building where the lecture is taking place, the result would satisfy clause (a) of the definition of optimal relevance, because (12) would be more relevant than anything the lecturer might have been saying at the time. Giora disagrees. She claims that (12) would be *less* relevant – though more important – than anything the lecturer might have been saying:

> "It is quite possible that the lecturer's current utterance, in the context of her talk, may achieve a greater amount of S&W relevance than the utterance in [(12)]. Nevertheless, there is no doubt that this more

---

[6] In a footnote, Giora suggests that the relevance-theoretic approach to context fails to take account of the role of mutual knowledge and the fact that speakers and hearers have to co-ordinate what they mutually know: "Mutual knowledge is not only a result of comprehension, but a prerequisite for it as well" (Giora, 1997, fn. 5). In fact, one of the earliest publications in relevance theory argued that mutual knowledge was *not* a prerequisite to comprehension, and in later work, we suggested that the notion of mutual knowledge should be replaced by a cognitively more plausible notion of mutual manifestness (see Sperber and Wilson, 1982; 1986/1995: 15–21, 38–46). The issue has been much discussed in the literature on relevance theory; for a summing up, see Garnham and Perner (1990); Sperber and Wilson (1990).

[7] For similar objections, see Morgan and Green (1987); for discussion, see Sperber and Wilson (1987: 746).

S&W relevant contribution will be abandoned for the sake of the new and less S&W relevant information." (Giora, 1997: 30)

In her view, our 'quantitative'[8] notion of relevance can shed no light on this example, and an explanation in terms of content is needed:

"Obviously, the attractiveness of [(12)] has nothing to do with its amount of uncostly contextual effects, but with its specific content, importance or quality. A contextual implication that has to do with life and death is worth a thousand less fatally laden." (Giora, 1997: 30)

These comments raises a specific question about the relevance of (12), and a more general one about the relation between relevance and importance or interest. First, the specific question. Would you find it relevant, in the middle of a lecture, to be told that the building you were sitting in was on fire? Would this information have any implications for you? Would you find it worth your effort to compute them at the time? If so, then (12) would be relevant enough to be worth your attention. In fact, in the situation described, it would be very relevant indeed.

Giora's belief that (12) is low in relevance is linked to her belief that "a contextual implication that has to do with life and death is worth a thousand less fatally laden". She fails to notice that on the relevance-theoretic account, a contextual implication that has to do with life and death will *yield* a thousand more. More generally, what makes information seem interesting or important is the fact that it can be productively processed; the more 'important' or 'interesting' the information, the greater will be the contextual effects, the more efficient the processing, and hence the greater the relevance.[9]

4. Coherence and optimal relevance

Giora's main aim is to show that, whatever its merits as a theory of comprehension, relevance theory cannot account for intuitions of coherence and degrees of coherence. As noted above, most coherence-based approaches aim to account for two types of intuition: (a) local intuitions about when two adjacent discourse segments are related (to each other or to a common topic); (b) global intuitions about

---

[8] In *Relevance*, pp. 129–132, we describe our definition of relevance as comparative rather than quantitative, and argue that although it might be interesting to develop a quantitative notion for computational purposes, it is the comparative rather than the quantitative notion that plays a role in human communication and cognition.

[9] I suspect that Giora's intuitions about differences in importance or interest are based largely on differences in level of attention. When someone tells you your house is on fire, your level of attention is bound to go up. More generally, the subjects that are intuitively 'interesting' or 'important' are those that lead to raised levels of attention. These alterations in level of attention are straightforwardly explainable in a relevance-based cognitive system, where processing effort is a function of time, on the one hand, and level of attention, on the other. An increase in level of attention makes it possible to compute more contextual effects in a given time. Hence, the more contextual effects are expected, the higher should be the level of attention. Giora's intuitions can be naturally accounted for in relevance-theoretic terms.

when a discourse 'hangs together' as a whole. Giora is more interested in global than local coherence, and, unlike many coherence theorists, she does not believe that local coherence is analysable in terms of a set of local coherence relations holding between adjacent segments in a text. However, she does acknowledge (fn. 4) that there are intuitions about the relatedness of adjacent segments. In the next section, I will look briefly at how relevance theory might account for such intuitions, before turning to the questions of global coherence that are Giora's main concern.

4.1. Local coherence and relevance

Giora uses examples (13) and (14) to illustrate her claim that a discourse may be 'S&W relevant' but nonetheless judged incoherent:

(13) Bill, who has thalassemia, is getting married to Susan, and 1967 was a great year for French wines.
(14) Bill, who has thalassemia, is getting married to Susan. Both he and Susan told me that 1967 was a great year for French wines.

In many circumstances, (13) and (14) would be indeed be unacceptable to some degree. Most coherence theorists would attribute their unacceptability to the fact that the two discourse segments are intuitively unrelated. Giora claims that relevance theory cannot explain their unacceptability: in her view, it predicts that they will be acceptable as long as each discourse segment is relevant in some context accessible to the hearer, whether or not they are intuitively related. She concludes that relevance theory must be supplemented by coherence theory if the full range of intuitions is to be explained.

I have already argued that the notion of 'S&W relevance' is inappropriate here, and should be replaced by the notion of optimal relevance and the criterion of consistency with the (communicative) principle of relevance. The genuine question raised by these examples is whether the criterion of consistency with the principle of relevance can explain why they would be unacceptable in many circumstances. I will argue that it can. I also want to make two more general points. First, the unacceptability of discourses such as (13)–(14) has more to do with wasted processing effort than with the fact that their discourse segments are unrelated. Second, intuitions about the relatedness of discourse segments can be straightforwardly captured using the machinery of relevance theory.

Notice, first, that there are circumstances in which (13) or (14) would be both acceptable and consistent with the principle of relevance. I will consider two types of case: in the first, their discourse segments will be intuitively unrelated; in the second, they are intuitively related.

Case A: Peter and Mary are catching up on the news of the day while clearing out a kitchen cupboard. Mary is about to tell Peter her news about Bill and Susan, when he holds up a bottle of 1967 French wine with a questioning look. She replies as in (13). In this case, each discourse segment would be relevant in some context accessible to Peter, but they would not be intuitively related.

In *Relevance*, we describe a similar situation:

"There is good reason to think that the memory of the deductive device is not the only short-term memory store available. Consider the fact that someone may divide his attention between two tasks: say, watching television and discussing family affairs at the same time. When this happens, it seems that he may be switching back and forth between two quite different contexts." (Sperber and Wilson, 1986/1995: 139)

(13) would be perfectly acceptable as a case of context switching. It would be consistent with the principle of relevance as long as Mary could reasonably have expected Peter to make the switch with ease. However, it would not be locally coherent, and the two discourse segments would be intuitively unrelated. We could explain this intuition by noting that the interpretation of the first segment makes no difference to the relevance of the second.[10]

Case B: Peter and Mary are catching up on the events of the day. Mary has some news about Bill and Susan, which she expresses as in (14). As the utterance proceeds, Peter accesses and uses the following contextual assumptions, in a way that Mary could manifestly have foreseen:

(15) a. People with thalassemia drink only red wine.
b. When people get married, it's usual to give a present.
c. A crate of wine is a suitable wedding present.
d. The best present is one that pleases the recipient.

Interpreting the utterance along these lines, Peter recovers the implicature that a crate of 1967 French red wine would be a good present to give to Bill and Susan.

In these circumstances, (14) would be acceptable and consistent with the principle of relevance. Its discourse segments would also be intuitively related. We could explain this intuition by noting that the interpretation of the first segment makes a difference to the relevance of the second: that is, different effects are achieved when the second segment is processed in the initial context, on the one hand, and when it is processed in the initial context supplemented by the interpretation of the first segment, on the other.

There are also several different types of case in which (13) and (14) would be unacceptable and inconsistent with the principle of relevance. I will consider two. In the first, the two discourse segments are intuitively related; in the second, they are not.

---

[10] This proposal is a refinement of an early relevance-theoretic attempt to capture the notion of *relevance of one remark to another* (Smith and Wilson, 1979: 177). That proposal failed to distinguish between contextual assumptions which contributed to relevance and those whose presence in the context was merely incidental. In *Relevance* (p. 139), we suggest that the machinery for making this distinction is needed on independent grounds. At the end of any discourse segment, it should therefore be possible to distinguish which assumptions have played a role in its interpretation and which have merely been incidental.

Case C: This is the same as Case B, except that Peter has access only to the contextual assumptions in (15b–d). As a result, the information that Bill has thalassemia will be irrelevant, but the two main discourse segments will be related as in Case B. On this interpretation, (14) will be unacceptable and inconsistent with the principle of relevance, despite the fact that its discourse segments are related.

Case D: This is similar to one that Giora considers in her paper. Peter and Mary are catching up on the news of the day. Peter tells Mary that their neighbour has just bought them a bottle of 1967 French wine. She replies as in (13), intending the first segment to be relevant in a context other than the one determined by Peter's immediately preceding remark, and the second to be a response to Peter.

In this case, each discourse segment would be relevant in some context accessible to Peter, but the two segments would be intuitively unrelated. According to Giora, the resulting discourse would be unacceptable in a way that relevance theory cannot explain. Let us assume that (13) would indeed be unacceptable (i.e. that it is not a straightforward case of context switching as in Case A above). Then its unacceptability can be explained along the following lines. A hearer with a presumption of optimal relevance should expect each segment of (13) to be relevant in the most accessible context. In the absence of context-switching cues, this will be the one created by the interpretation of the immediately preceding segment. If Mary did not intend Peter to use this context, she could have saved him some effort by reformulating her utterance as in (16):

(16) 1967 was a great year for French wine. Incidentally, Bill, who has thalassemia, is getting married to Susan.

Thus (13), on the interpretation intended by Giora, is inconsistent with the principle of relevance, and its unacceptability is explained.

The analysis of these four cases suggests that the criterion of consistency with the principle of relevance can shed more light on intuitions of acceptability and unacceptability than the appeal to local coherence. What makes a discourse like (13) or (14) unacceptable in the circumstances described is the manifest waste of effort spent in looking for relevance in the wrong direction, or failing to find it at all. The resulting interpretations will be inconsistent with the principle of relevance, whether or not the discourse segments are related. By the same token, what makes a discourse like (13) or (14) acceptable in the circumstances described is the fact that it has an interpretation consistent with the principle of relevance, whether or not the discourse segments are related.

4.2. Global coherence and relevance

Giora's own account of discourse coherence can be summarised as follows (Giora, 1997: 22–23):

(17) An informative discourse is well-formed iff it:
   a. conforms to the Relevance Requirement in that all its propositions are conceived of as related to a discourse-topic proposition;

b. conforms to the Graded Informativeness Condition, which requires that each proposition be more (or at least not less) informative than the one that precedes it in relation to the discourse topic; and
c. marks any deviation from Relevance and Graded Informativeness by an explicit marker, e.g. 'by the way', 'after all'.

This is not the place for a detailed discussion of Giora's proposal; I will merely comment briefly on each of the clauses in her definition, and sketch an alternative relevance-theoretic account.

4.2.1. The Relevance Requirement

The coherence-based literature offers two main approaches to global coherence. On the first, what makes a discourse hang together is a hierarchical structure of coherence relations (e.g. Hobbs, 1990; Hovy, 1990). On the second, which Giora adopts, what makes a discourse hang together is a hierarchical structure of discourse topics. The notion of discourse topic has considerable intuitive appeal. A well-planned academic article, with its structure of sections, sub-sections and sub-sub-sections, seems to be just what ought to be expected on a topic-based approach. However, Dan Sperber and I have argued that the notion of topic-relevance is derivative, and can be dispensed with on a relevance-theoretic account.

It is widely accepted that the function of a discourse topic is to provide access to contextual information required for the comprehension of the associated text or discourse. In *Relevance* (p. 216–217), we argue that what is crucial to comprehension is the contextual information rather than the discourse topic. Thus, a range of experiments which have been seen as providing decisive evidence about the need for discourse topics can be reanalysed in relevance-theoretic terms. To take just one example: Giora cites some experiments by Bransford and Johnson (1972) which seem to show that certain texts are incomprehensible unless their discourse topic is explicitly stated (e.g. in a title). In Sperber and Wilson (1987: 742), we give two reasons for thinking that what makes these passages incomprehensible may be not so much the absence of an explicit discourse topic as the inaccessibility of the contextual information to which this topic would have given access. In the first place, even where there is an explicitly stated topic, such a text would remain incomprehensible if the required contextual information were lacking. In the second place, even where there is no explicitly stated topic, such a text should become comprehensible if the required contextual information is made accessible enough in some other way.

In fact, Giora's views on the need for explicit statement of discourse topics are not quite clear. In analysing the following example, she claims that a discourse topic must be explicitly stated:

"At one point, it was obvious to my students and to me that I would start my lecture stating that I was leaving the university soon ... However, even though I knew they knew I was going to discuss this topic upon the beginning of the lecture ..., I could not ... start in medias res. I had to start by stating my discourse topic ... : 'As you well know, I am leaving soon, etc.'" (Giora, 1997: 24).

In her view, the announcement that she was about to leave presents problems for relevance theory, because it is 'S&W irrelevant' but nonetheless coherent and acceptable. She claims that, by contrast, its acceptability can be accounted for in terms of her Relevance Requirement, which "requires that the discourse-topic be stated and presented in discourse-initial position, and then elaborated on" (1997: 24). Actually, the Relevance Requirement (see (17) above) is much weaker, stating merely that the discourse topic should 'preferably' be made explicit, and is 'better' stated than assumed. Indeed, discussing the Relevance Requirement elsewhere in her paper, she comments:

"Note that the discourse topic need not be made explicit in the text. It is enough that it can be generated." (Giora, 1997: fn. 7)

I will try to show that relevance theory can shed more light on this example than her Relevance Requirement can.

Giora believes that her announcement would be irrelevant (in Sperber and Wilson's sense) in the situation described. In fact, it would not be irrelevant at all. One way of achieving relevance is by confirming expectations. In the situation described, the students have two expectations: first, that their professor is going to leave the university, and second, that she is going to tell them she is leaving. Her announcement would achieve relevance by confirming the first expectation and verifying the second. It would also contribute to relevance in two further ways. In its function as 'discourse topic', it would provide easy access to a chunk of contextual information required for the comprehension of later utterances, allowing the discourse as a whole to achieve optimal relevance. And it would make not merely manifest but mutually manifest the fact that she was leaving. As we note in *Relevance*:

"Mere informing alters the cognitive environment of the audience. Communication alters the mutual cognitive environment of the audience and communicator. Mutual manifestness may be of little cognitive importance, but it is of crucial social importance. A change in the mutual cognitive environment of two people is a change in their posibilities of interaction (and, in particular, in their possibilities of future communication)." (Sperber and Wilson, 1986/1995: 61–62)

Thus, relevance theory seems to shed more light on this example than Giora's own account.

4.2.2. *The Graded Informativeness Condition*

According to the Graded Informativeness Condition, each proposition in a given discourse segment must be informative enough to eliminate some existing alternatives; however, it must not be so informative as to violate the Relevance Requirement. Giora uses (18) as an example of a text in which both conditions are satisfied:

(18) It has often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter unexpectedly came upon another which was far more important than the one he was looking into. Penicillin is a result of such a discovery.

She then considers two possible continuations: (19), in which a new utterance is added within the same paragraph, and (20), in which it starts a new paragraph of its own:

(19) It has often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter unexpectedly came upon another which was far more important than the one he was looking into. Penicillin is a result of such a discovery. Penicillin was accidentally discovered by Fleming in 1928 ...

(20) It has often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter unexpectedly came upon another which was far more important than the one he was looking into. Penicillin is a result of such a discovery.
Penicillin was accidentally discovered by Fleming in 1928 ...

Twenty experimental subjects who read these passages judged (20) more coherent or appropriate than (19). Giora believes that these results are explained by an interaction between her Relevance Requirement and her Graded Informativeness Condition: the final sentence adds too much new information to be included within the same paragraph, as in (19), and must start a new paragraph of its own. She claims that relevance theory sheds no light on these examples, and more generally on the existence and placement of paragraph breaks.

In fact, several studies of paragraph structure (e.g. Stark, 1988; Hofmann, 1989) suggest a different explanation for her experimental results. One of the strongest cues to paragraph boundaries is the use of a full NP where a pronoun would do, as with the repetition of 'Penicillin' in the final sentence of (19) and (20) above. It would be interesting to test whether the same results would be obtained with 'It' substituted for 'Penicillin'. Stark's experiments suggest that they would not; by contrast, since pronominalisation should not affect informativeness or topic-relevance, Giora predicts that they would.

Stark's experiments also show that in the absence of cues such as those in (19) and (20) above, the removal or relocation of existing paragraph breaks does not affect acceptability ratings or judgements of coherence. It does, however, affect judgements about what information is important, and in a rather interesting way. When the author's original paragraph markings were left intact, subjects judged the first sentence of a paragraph more important than they did when the paragraph markings were removed; however, misplaced boundaries had no effect on ratings of importance. This shows that paragraph marking does affect comprehension, but in a way that is relatively constrained. What light can relevance theory shed on these results?

Relevance theory fits well with the view of Chafe (1987: 48) that discourse is best approached in terms of process rather than structure:

"It is more rewarding, I think, to interpret a piece of discourse in terms of cognitive processes dynamically unfolding through time than to analyse it as a static string of words and sentences."

On Chafe's account, paragraph breaks coincide with wholesale shifts of information from active to semi-active state. In more familiar terms, they coincide with major reorganisations of context. In a well-organised text, each utterance within a given paragraph should be interpretable in a context partly determined by the interpretation of its immediate predecessor. The processes that go on around paragraph breaks (anaphoric processes, intonation changes, pauses, slowed reading speeds, etc.) save the addressee wasted effort by alerting him to the fact that a switch in contexts is about to take place. It seems to me that Giora is quite sympathetic to accounts along these general lines, and that our approaches are not incompatible.

4.2.3. Deviations from Relevance and Graded Informativeness

Clause (c) of Giora's definition of coherence deals with violations of clauses (a) and (b). It allows for deviations from the Relevance Requirement and the Graded Informativeness Condition, as long as the deviations are explicitly marked, e.g. by use of the markers 'by the way' or 'after all'. Here, too, I want to argue that relevance theory offers better explanations than Giora's coherence-based account.

Consider the exchange in (21):

(21) a. *Peter*: What time is it?
b. *Mary*: You've dropped your wallet.

(21) violates the Relevance Requirement, and in Giora's framework it would be ill-formed. The fact that it is clearly both comprehensible and acceptable is not seen as problematic; discussing a similar example, Giora comments:

"On various occasions, proponents of S&W relevance theory ... made the claim that the frequent occurrence of (relatively) incoherent but S&W relevant discourses [e.g. (21)] attests that speakers and hearers are not necessarily constrained by the search for coherence. This claim is not problematic for a coherence-based theory. A coherence-based theory endeavors to account for and make explicit speakers' intuitions as to the well-formedness of discourses. It does not assume that coherence is the only principle that governs human communication." (Giora, 1997: 28)

This amounts to the claim that (21) is acceptable but ill-formed.

Now consider the exchange in (22):

(22) a. *Peter*: What's the time?
b. *Mary*: By the way, you've dropped your wallet.

(22) again violates the Relevance Requirement. Again, the result will be both comprehensible and acceptable, but this time it will also be well-formed in Giora's framework, because the deviation is explicitly marked. The resulting analysis, with its sharp distinction between 'well-formed' and 'ill-formed' deviations, seems to me to bring into question the whole coherence-based approach.

One problem with clause (c) of Giora's definition is that it places no constraint on the content or number of deviations allowed. Thus, in her framework, (23) and (24) will be well-formed:

(23) a. *Peter*: What's the time?
b. *Mary:* By the way, Tutankhamun ate my dog.

(24) Today is Tuesday. Incidentally, I can dance the tango. By the way, there are tadpoles in my village pond. Anyway, how many relevance theorists does it take to change a light-bulb? ...

A theory that rejects (21) while accepting (23) and (24) seems to me to leave something to be desired. Giora does suggest that the content of deviations may be constrained by considerations of 'interest' or 'importance'. By the same token, they will be constrained by considerations of optimal relevance, which will also limit the number of deviations considered acceptable in a given text.

Consider, next, the rationale behind clause (c). What is the function of explicit markers such as 'by the way', 'anyway', etc., and why should they make a deviation more acceptable? In the relevance-theoretic framework, they could be seen as saving the addressee wasted effort by indicating that a change of context was coming up. However, the same function can be performed by non-linguistic means. For example, in (21) above, if Mary thought that Peter might fail to make the expected switch in context, she might have pointed, touched Peter's arm, or turned her head in the direction of the dropped wallet. The result would have been both acceptable and consistent with the principle of relevance despite the absence of explicit linguistic cues. In fact, in many circumstances the switch required in (21) would be made without any cues at all. Suppose that as, Peter was speaking to Mary, it suddenly struck him that his wallet might have gone. Then there would be no need for cues to context-switching: the required assumptions would have been activated before Mary spoke.

Notice, too, that the presence of explicit linguistic markers does not always reduce the possible lines of interpretation, as Giora seems to suppose. Consider (25):

(25) a. *Peter*: What did Susan say?
b. *Mary*: You've dropped your wallet.

(25b) has two interpretations: a locally coherent one on which Mary is reporting that Susan said 'You've dropped your wallet', and a non-coherent one, similar to those in (21a) above. As we have seen, Giora regards the non-coherent interpretation of (25) as ill-formed because it deviates from the Relevance Requirement.[11] According to clause (c) of her definition, she should regard (26) as well-formed because the deviation is explicitly marked:

(26) a. *Peter*: What did Susan say?
b. *Mary*: By the way, you've dropped your wallet.

---

[11] Giora's analysis of this example makes two disputable claims which I have already discussed: (a) that the less coherent interpretation is also less relevant; hence (b) that the hearer's willingness to attend to this interpretation should be explained by an appeal to 'importance' or 'interest' rather than relevance.

Notice, though, that despite the explicit marking, (26b) is just as ambiguous as (25b). So why should (26) be treated as well-formed when (25) is not? It seems to me that the presence or absence of explicit linguistic markers is much too crude a diagnostic: what best explains the acceptability or unacceptability of 'deviations' of the type discussed above is the notion of optimal relevance and the criterion of consistency with the principle of relevance.[12]

References

Berg, Jonathan, 1991. The relevant relevance. Journal of Pragmatics 16: 411–425.

Blakemore, Diane, 1988. The organisation of discourse. In: Frederick Newmeyer, ed., Language: The sociocultural context, Vol. 4: 229–250. Cambridge: Cambridge University Press.

Blass, Regina, 1986. Cohesion, coherence and relevance. Notes on Linguistics 34: 41–64.

Blass, Regina, 1990. Relevance relations in discourse: A study with special reference to Sissala. Cambridge: Cambridge University Press.

Bransford, John and Marcia Johnson, 1972. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior 11: 717–726.

Brown, Gillian and George Yule, 1983. Discourse analysis. Cambridge: Cambridge University Press.

Chafe, Wallace, 1987. Cognitive constraints on information flow. In: R. Tomlin, ed., Coherence and grounding in discourse, 21–51. Amsterdam: Benjamins.

Garnham, Alan and Josef Perner, 1990. Does manifestness solve problems of mutuality? Behavioral and Brain Sciences 13(1): 178–179.

Giora, Rachel, 1985a. Towards a theory of coherence. Poetics Today 6(4): 699–716.

Giora, Rachel, 1985b. A text-based analysis of non-narrative discourse. Theoretical Linguistics 12(2/3): 115–135.

Giora, Rachel, 1988. On the informativeness requirement. Journal of Pragmatics 12(5/6): 547–565. Reprinted in Asa Kasher, ed., 1989, Cognitive aspects of language use, 63–96. Amsterdam: North-Holland.

Giora, Rachel, 1996. Language comprehension as structure building. Journal of Pragmatics 26: 417–436.

Giora, Rachel, 1997. Discourse coherence and theory of relevance: Stumbling blocks in search of a unified theory. Journal of Pragmatics 27: 17–34.

Hobbs, Jerry, 1978. Why is discourse coherent? Menlo Park, CA: SRI International. Technical Note 176.

Hobbs, Jerry, 1990. Topic drift. In: Bruce Dorval, ed., Conversational organisation and its development, 3–22. Norwood, NJ: Ablex.

Hobbs, Jerry, M. Stickel and R. Martin, forthcoming. Interpretation as abduction. To appear in Artificial Intelligence.

Hofmann, Thomas, 1989. Paragraphs and anaphora. Journal of Pragmatics 13: 239–250.

Hovy, Ed, 1990. Parsimonious and profligate approaches to the question of discourse structure relations. In: Proceedings of the 5th International Workshop on Natural Language Generation, 128–136.

Mann, William and Sandra Thompson, 1986. Relational propositions in discourse. Discourse Processes 9: 57–90.

Morgan, Jerry and Georgia Green, 1987. On the search for relevance. Behavioral and Brain Sciences 10(4): 726–727.

Reinhart, Tanya, 1981. Pragmatics and linguistics: An analysis of sentence topics. Philosophica 27: 53–94.

Samet, J. and Roger Schank, 1984. Coherence and connectivity. Linguistics and Philosophy 7: 57–82.

[12] For further comparison of relevance-based and coherence-based approaches to discourse, see Blakemore (1988); Blass (1986, 1990); Unger (1996).

Smith, Neil and Deirdre Wilson, 1979. Modern linguistics: The results of Chomsky's revolution. Harmondsworth: Penguin.

Smith, Neil and Deirdre Wilson, 1992. Introduction to the special issue on relevance theory. Lingua 87: 1–10.

Sperber, Dan, 1994. Understanding verbal understanding. In: Jean Khalfa, ed., What is intelligence, 179–198. Cambridge: Cambridge University Press.

Sperber, Dan and Deirdre Wilson, 1982. Mutual knowledge and relevance in theories of comprehension. In: Neil Smith, ed., Mutual knowledge, 61–131. London: Academic Press.

Sperber, Dan and Deirdre Wilson, 1986/1995. Relevance: Communication and cognition. Oxford: Blackwell.

Sperber, Dan and Deirdre Wilson, 1987. Presumptions of relevance. Behavioral and Brain Sciences 10(4): 736–754.

Sperber, Dan and Deirdre Wilson, 1990. Reply to Garnham and Perner. Behavioral and Brain Sciences 13(1): 183–184.

Stark, Heather, 1988. What do paragraph markings do? Discourse processes 11: 275–303.

Unger, Christoph, 1996. The scope of discourse connectives: Implications for discourse organisation. Journal of Linguistics 32: 403–438.