

# Précis of *Relevance: Communication and Cognition*

**Dan Sperber**

*C.N.R.S., 15 Quai Anatole France, 75007 Paris, France<sup>1</sup>*

**Deirdre Wilson**

*Department of Linguistics, University College London, London WC1E 6BT, England*

**Abstract:** In *Relevance: Communication and Cognition*, we outline a new approach to the study of human communication, one based on a general view of human cognition. Attention and thought processes, we argue, automatically turn toward information that seems relevant: that is, capable of yielding cognitive effects – the more, and the more economically, the greater the relevance. We analyse both the nature of cognitive effects and the inferential processes by which they are derived.

Communication can be achieved by two different means: by encoding and decoding messages or by providing evidence for an intended inference about the communicator's informative intention. Verbal communication, we argue, exploits both types of process. The linguistic meaning of an utterance, recovered by specialised decoding processes, serves as the input to unspecialised central inferential processes by which the speaker's intentions are recognised.

Fundamental to our account of inferential communication is the fact that to communicate is to claim someone's attention, and hence to imply that the information communicated is relevant. We call this idea, that communicated information comes with a guarantee of relevance, the *principle of relevance*. We show that every utterance has at most a single interpretation consistent with the principle of relevance, which is thus enough on its own to account for the interaction of linguistic meaning with contextual factors in disambiguation, reference assignment, the recovery of implicatures, the interpretation of metaphor and irony, the recovery of illocutionary force, and other linguistically underdetermined aspects of utterance interpretation.

**Keywords:** attention; communication; comprehension; implicature; irony, metaphor; nondemonstrative inference; pragmatics; relevance; speech acts; style; thinking

In *Relevance: Communication and Cognition* (Sperber & Wilson 1986a, henceforth *Relevance*), we present a new approach to the study of human communication. This approach, outlined in the first chapter, is grounded in a general view of human cognition developed in the second and third chapters. Human attention and thought, we argue, automatically turn toward information which seems relevant: To communicate is to claim someone's attention, hence to communicate is to imply that the information communicated is relevant. We call this thesis the *principle of relevance*, and show in the fourth chapter how it is enough on its own to account for the interaction of linguistic meaning with contextual factors in utterance interpretation.

In this précis, we will follow the general plan of the book. However, we have had to leave out several steps in the argumentation, many side issues, most examples, almost all discussion of other approaches, and all traces of wit.

## 1. Communication

**1.1. The code model and its limits.** Communication is a process involving two information-processing devices. One device modifies the physical environment of the other. As a result, the second device constructs representations similar to representations already stored in the

first device. Oral communication, for instance, is a modification by the speaker of the hearer's acoustic environment, as a result of which the hearer entertains thoughts similar to the speaker's own.

The question is: How can a physical stimulus bring about the required similarity of representations when there is no similarity whatsoever between the stimulus (e.g., sound patterns) on the one hand and the representations (e.g., human thoughts) it brings into correspondence on the other? From Aristotle through to modern semiotics, all theories of communication were based on a single model, which we call the *code model*. A *code* is a system which pairs internal messages with external signals, thus enabling two information-processing devices (organisms or machines) to communicate.

Linguistic utterances – the most important means of human communication – do succeed in communicating thoughts; the hypothesis that utterances are signals that encode thoughts seems to explain this fact. However, it is descriptively inadequate: Comprehension involves more than the decoding of a linguistic signal. Although a language can be seen as a code which pairs phonetic and semantic representations of sentences, much recent work in psycholinguistics, pragmatics, and the philosophy of language<sup>2</sup> shows that there is a gap between the semantic representations of sentences and the thoughts actually communicated by utterances. This gap is filled not by more coding, but by inference.

The study of the semantic representation of sentences belongs to grammar; the study of the interpretation of utterances belongs to what is now known as *pragmatics*. Among its tasks, pragmatics must explain how hearers resolve ambiguities, complete elliptical or otherwise semantically incomplete sentences, identify intended references, identify illocutionary force, recognize tropes, and recover implicit import. These are some of the ways in which the context-independent semantic representation of a sentence falls short of determining the interpretation of an utterance of that sentence in context.

To justify the code model of verbal communication, it would have to be shown that the interpretation of utterances in context can be accounted for by adding an extra pragmatic level of decoding to the linguistic level provided by the grammar. Much recent work in pragmatics has assumed that this can be done.<sup>3</sup> At the programmatic level, pragmatics has been described, on the analogy of phonology, syntax, and semantics, as a code-like mental device underlying a distinct level of linguistic ability. In practice, however, most pragmatics have described comprehension as an inferential process.

Inferential and decoding processes are quite different. An *inferential process* takes a set of premises as input and yields as output a set of conclusions which follow logically from, or are at least warranted by, the premises. A *decoding process* takes a signal as input and yields as output a message associated with the signal by an underlying code. In general, conclusions are not associated with their premises by a code, and signals do not warrant the messages they convey. Does it follow that pragmatics who hold to the code model but describe comprehension in inferential terms are being inconsistent? Not necessarily: It is formally conceivable that a decoding process should contain an inferential process as a subpart. However, for this to be possible, speaker and hearer must use not only the same language but also the same set of premises, because what makes the code model explanatory is that symmetrical operations are performed at the emitting and receiving ends.

The set of premises used in interpreting an utterance constitutes what is generally known (see Gazdar 1979; Johnson-Laird 1983) as the *context*. A context is a psychological construct, a subset of the hearer's assumptions about the world. Each new utterance, though drawing on the same grammar and the same inferential abilities as previous utterances, requires a rather different context (if only because the interpretation of the previous utterance has become part of the context). A central problem for pragmatic theory is to describe how the hearer constructs a new context for every new utterance.

For code theorists, the context used by the hearer should always be identical to the one envisaged by the speaker. Can this condition be met? Because any two people are sure to share at least a few assumptions about the world, they might be expected to use only these shared assumptions. However, this cannot be the whole answer in that it immediately raises a new question: How are the speaker and hearer to distinguish the assumptions they share from those they do not? For that, they must make second-order assumptions about which first-order assumptions they share; but then they had better make sure that they share these second-order assumptions,

which calls for third-order assumptions, and so on indefinitely. Assumptions or knowledge of this infinitely regressive sort was first identified by Lewis (1969) as "common knowledge" and by Schiffer (1972) as "mutual knowledge."

Within the framework of the code model, mutual knowledge is a necessity. However, pragmatics have offered no independent support for the claim that individuals engaging in verbal communication can and do distinguish mutual from nonmutual knowledge. In *Relevance*, we present several arguments to show that the mutual knowledge hypothesis is psychologically implausible. We therefore reject the code model of verbal communication that implies it.<sup>4</sup>

**1.2. The inferential model.** In 1957, Paul Grice published an article, "Meaning," which has been the object of a great many controversies, interpretations, and revisions.<sup>5</sup> In it Grice analysed what it is for an individual to mean something by an utterance in terms of intentions and the recognition of intentions and tried to extend this analysis of "speaker's meaning" into such areas of traditional semantic concern as the analysis of "sentence meaning" and "word meaning." Grice's analysis provides the point of departure for a new model of communication, the *inferential model*, and this is how we use it in *Relevance*. We look in detail at Grice's own proposal, as well as at some of the objections that have been raised and some of the reformulations that have been proposed, notably by Strawson (1964), Searle (1969), and Schiffer (1972). Here we will give a short informal account of the basic idea before outlining some developments of our own.

Suppose that Mary intends to inform Peter of the fact that she has a sore throat. All she has to do is let Peter hear her hoarse voice, thus providing him with salient and conclusive evidence that she has a sore throat. Suppose now that Mary intends, on June 2, to inform Peter that she had a sore throat on the previous Christmas Eve. This time she is unlikely to be able to produce *direct* evidence of her past sore throat. What she can do though is give him direct evidence, not of her past sore throat, but of her present intention to inform him of it. She may do this, for instance, by saying "I had a sore throat on Christmas Eve," or by nodding when he asks her if she did. Mary's utterance or nod is directly caused by her present intention to inform Peter of her past sore throat and is therefore direct evidence of this intention. Suppose now that Peter assumes that Mary is sincere and well-informed. Then the fact that she intends to inform him that she had a sore throat on that date provides indirect but nevertheless conclusive evidence that she had one. Mary's intention to inform Peter of her past sore throat is fulfilled by making him recognise her intention.

This example shows that information can be conveyed in two different ways. One way is to provide direct evidence for it. This should not in itself be regarded as a form of communication: Any state of affairs provides direct evidence for a variety of assumptions without necessarily *communicating* those assumptions in any interesting sense. Another way of conveying information is to provide direct evidence of one's intention to convey it. The first method can be used only with information for which direct evidence can be provided. The second

method can be used with any information at all, as long as direct evidence of the communicator's intentions can be provided. This second method is clearly a form of communication; it can be called *inferential communication* because the audience infers the communicator's intention from evidence provided for this precise purpose.

A communicator intentionally engaging in inferential communication perceptibly modifies the physical environment of her<sup>6</sup> audience – that is, she produces a stimulus. She does so with two characteristic intentions: the *informative intention*, to inform the audience of something, and the *communicative intention*, to inform the audience of her informative intention. Note that the communicative intention is itself a second-order informative intention.

This description of communication in terms of intentions and inferences is, in a way, commonsensical. As speakers, we intend our hearers to recognise our intention to inform them of some state of affairs. As hearers, we try to recognise what it is that the speaker intends to inform us of. The idea that communication exploits the well-known ability of humans to attribute intentions to each other should appeal to cognitive and social psychologists. To justify this appeal, however, what is needed is not merely a descriptive account, but a genuine explanation of communicative success.

How are informative intentions recognised? The key to an answer is again suggested by Grice (1975a, 1978), who argues that a rational communicator tries to meet certain general standards. Grice describes these standards as a "cooperative principle" and nine associated "maxims." From knowledge of these standards, observation of the communicator's behaviour, and the context, the audience can normally infer the communicator's informative intention. In *Relevance*, we discuss in detail both Grice's proposal and its elaborations by others. Here a brief illustration must suffice.

Consider the following dialogue:

(1) *Peter*: Do you want some coffee?

*Mary*: Coffee would keep me awake.

Unless some further assumptions are made, Mary's answer fails to satisfy one of Grice's maxims: "Be relevant." However, Peter should take for granted that Mary is not flouting the maxim; he can justify this assumption by assuming, further, that she intends him to infer from her answer that she does not want to stay awake and hence does not want any coffee. These contextually inferred assumptions, recovered by reference to the cooperative principle and maxims, are what Grice calls *implicatures* of her utterance. Such implicatures are communicated not by coding, but by providing evidence of the fact that the speaker intends to convey them.

Although more systematic than the reconstructions that can be elicited from unsophisticated speakers, the analyses of implicature proposed by Grice and his followers have shared with these reconstructions the defect of being almost entirely *ex post facto*. Given that an utterance in context is found to carry particular implicatures, what both the hearer and the pragmatician can do is to show how, in intuitive terms, an argument based on the context, the utterance, and general expectations about the behaviour of speakers justifies the particular

interpretation chosen. What they fail to show is that, on the same basis, an equally well-formed argument could not have been given for a quite different and in fact implausible interpretation.

Grice's idea that the very act of communicating creates expectations which it then exploits provides a starting point. Beyond that, the inferential model needs radical reworking in order to become truly explanatory. A psychologically realistic answer must be given to such basic questions as these: What shared information is exploited in communication? What forms of inference are used? What is relevance and how is it achieved? What role does the search for relevance play in communication?

### 1.3. Cognitive environments and mutual manifestness.

In analysing the nature of the shared information involved in communication, we introduce the notion of a cognitive environment (analogous, at a conceptual level, to notions of visual or acoustic environment at a perceptual level):

A *cognitive environment* of an individual is a set of facts that are manifest to him.

A fact is *manifest* to an individual at a given time if, and only if, the individual is capable at that time of representing it mentally and accepting its representation as true or probably true.

To be manifest, then, is to be perceptible or inferable. An individual's total cognitive environment consists not only of all the facts that he is aware of, but of all the facts that he is capable of becoming aware of at that time and place. Manifestness so defined is a property not only of facts but, more generally, of true or false assumptions. It is a relative property: Facts and assumptions can be more or less strongly manifest. Because *manifest* is weaker than *known* or *assumed*, a notion of mutual manifestness can be developed that does not suffer from the same psychological implausibility as mutual knowledge.

The same facts and assumptions may be manifest in the cognitive environments of several people. In that case, these cognitive environments intersect, and their intersection is a cognitive environment that the people in question share. One thing that can be manifest in a shared cognitive environment is a characterization of those who have access to it. For instance, every Freemason has access to a number of secret assumptions, which include the assumption that all Freemasons have access to these same secret assumptions. In other words, all Freemasons share a cognitive environment that contains the assumption that all Freemasons share this environment. Any shared cognitive environment in which it is manifest which people share it is what we call a *mutual cognitive environment*. For every manifest assumption, in a mutual cognitive environment, the fact that it is manifest to the people who share the environment is itself manifest. In a mutual cognitive environment, therefore, every manifest assumption is *mutually manifest*.

If a cognitive environment is merely a set of assumptions that an individual is capable of mentally representing and accepting as true, the question becomes: Which of these assumptions will the individual actually make? This question is of interest not only to the psychologist, but also to every ordinary communicator who wants to modify the thoughts of her audience but can directly affect only its cognitive environment.



**1.4. Relevance and ostension.** Most discussions of information processing, whether in experimental psychology or in artificial intelligence, have been concerned with the achievement of preset goals. However, many cognitive tasks consist not in reaching a fixed state, but in increasing the value of some parameters. Human cognition as a whole is a case in point: It is aimed at improving the quantity, quality, and organization of the individual's knowledge. To achieve this goal as efficiently as possible, the individual must at each moment try to allocate his processing resources to the most *relevant* information: that is, as we will shortly show, information likely to bring about the greatest improvement of knowledge at the smallest processing cost. Our claim is that this is done automatically and that an individual's particular cognitive goal at a given time is always consistent with the more general goal of maximising the relevance of the information processed.

Human cognition is relevance-oriented. As a result, and to the extent that one knows the cognitive environment of an individual – which one does when the environment is mutual – one can infer which assumptions he is actually likely to entertain and how a change in that environment might affect his train of thoughts. This makes it possible to affect people's thoughts in a partly predictable way by modifying their cognitive environment.

Peter and Mary are sitting on a park bench. He points in a direction where she had not so far noticed anything in particular. This time, she takes a closer look and sees their acquaintance Julius in the distance, sitting on the grass. In other words, as a result of Peter's behaviour, the presence of Julius, which was weakly manifest in Mary's cognitive environment, has become more manifest, to the point of being actually noticed. Moreover, it has become manifest that Peter had himself noticed Julius and intended her to notice him too. Such behaviour – which makes manifest an intention to make something manifest – we call *ostension*.

How does ostension work? For instance, how does Mary discover, when Peter points in a certain direction, which of the many phenomena visible in that direction he intended her to notice? Any request for attention, and hence any act of ostension, conveys a presumption of relevance; it does so because attention goes only to what is presumed relevant. By pointing, Peter conveys to Mary that by paying attention she will gain some relevant information. This makes it reasonable for her to pay more attention than she had before; discovering the presence of Julius, she may reasonably assume that Peter thought it would be relevant to her, and, moreover, that this was part of his reason for bringing it to her attention.

Ostension provides two layers of information to be picked up: The first consists of the information that has been pointed out; the second consists of the information that the first layer of information was intentionally pointed out. In our example, the first basic layer of information – Julius's presence – was already manifest and is merely made more manifest by the ostension. In other cases, all the evidence displayed in an act of ostension bears on the agent's intentions and on nothing else directly. In these cases, it is only by discovering the agent's intentions that the audience can also discover, indirectly, the basic information that the agent intended to make manifest. As

we show in *Relevance*, there is a continuum of cases between those that provide full direct evidence for the basic information made manifest (i.e., they "show something") and those that provide only indirect evidence (e.g., by "saying something"). We argue that inferential communication and ostension are one and the same process, but seen from two different points of view: that of the communicator who is involved in ostension and that of the audience who is involved in inference.

**1.5. Ostensive–inferential communication.** Most accounts of communication take "saying that" as their paradigm case, and assume that the communicator's intention is to induce certain specific thoughts in an audience. We want to suggest that the communicator's informative intention is better described as an intention to modify directly not the thoughts but the cognitive environment of the audience, with only partly foreseeable effects on the audience's actual thoughts. We therefore reformulate the notion of an informative intention:

*Informative intention:* the intention to make manifest or more manifest to the audience a certain set of assumptions.

Why should someone who has an informative intention bother to make it known to the audience? In other words, what are the reasons for engaging in ostensive communication? So far we have discussed only one of these reasons: Making one's informative intention known is often the best way, or the only way, of fulfilling it. There is another major reason for engaging in ostensive communication. Mere informing alters the cognitive environment of the audience. True communication is "overt" in Strawson's sense (Strawson 1964); in our terms, it takes place in the mutual cognitive environment of the audience and communicator. Mutual manifestness may be of little cognitive importance, but it is of crucial social or interpersonal importance. A change in the mutual cognitive environment of two people is a change in their possibilities of interaction (and, in particular, in their possibilities of further communication). This is why we redefine the communicative intention as follows:

*Communicative intention:* the intention to make mutually manifest to audience and communicator the communicator's informative intention.

Ostensive–inferential communication, which, incidentally, need not be intentional, can itself be defined as follows:

*Ostensive–inferential communication:* The communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a certain set of assumptions.

Instead of treating an assumption as either communicated or not communicated, we have a set of assumptions which, as a result of communication, become manifest or more manifest to varying degrees. We can think of communication itself, then, as a matter of degree. When the communicator makes strongly manifest her informative intention to make some particular assumption strongly manifest, then that assumption is strongly communicated. An example would be answering a clear "Yes," when asked "Will you take John So-and-so as your lawful

wedded husband?" When the communicator's intention is to increase simultaneously the manifestness of a wide range of assumptions, so that her intention concerning each of these assumptions is itself weakly manifest, then each of them is weakly communicated. An example would be sniffing ecstatically and ostensibly at the fresh seaside air. Often in human interaction weak communication is found sufficient or even preferable to the stronger forms.

Nonverbal communication is often of the weaker kind. One of the advantages of verbal communication is that it allows the strongest possible form of communication to take place; it enables the hearer to pin down the speaker's intentions about the explicit content of her utterance to a single, strongly manifest candidate, with no alternative worth considering at all. On the other hand, what is implicitly conveyed in verbal communication is generally weakly communicated. Because all communication has been seen as strong communication, descriptions of nonverbal communication have been marred by spurious attributions of definite meaning; and in the case of verbal communication, the vagueness of most implicatures and of nonliteral forms of expression has been idealised away. Our approach, we believe, provides a way of giving a precise description and explanation of the weaker effects of communication.

We began this section by asking how communication is possible. Our answer is that it is possible in at least two very different ways: by means of a code shared by communicator and audience, or by means of ostensive stimuli providing the audience with evidence from which the communicator's informative intention can be inferred. We argue against upgrading either model to the status of a general theory of communication. It is particularly important to keep the distinction between the two modes of communication in mind when it comes to describing how they can combine, as they do in human verbal communication.

## 2. Inference

In the second chapter of *Relevance*, we outline a model of the main inferential abilities involved in verbal comprehension. This model is concerned with only one type of inferential process – that of nondemonstrative inference from assumptions to assumptions, which, we claim, takes place automatically and unconsciously during comprehension. We do not discuss conscious reasoning, which sometimes plays a role in comprehension; we merely suggest how unconscious inference may be exploited in conscious reasoning.

In presenting this model of inference, our aims are twofold. First, we claim that the general notion of relevance is instantiated differently in each particular inferential system. By describing one system, however sketchily, we make it possible to give a detailed example of such an instantiation. If future research shows that human inferential abilities not only are much more complex and varied than our model (which of course they are), but also radically different from that model, this illustrative purpose would still be served.

Our second aim is to show how the study of inferential comprehension may shed light on central thought processes. We accept in broad outline Jerry Fodor's view of

the modularity of mind (Fodor 1983; 1985); like him, we see linguistic decoding as modular. Unlike him, however, we see the inferential tier of verbal comprehension as the application of unspecialised central thought processes to the output of the linguistic decoding module. We argue that verbal comprehension is more typical of central processes, and much more amenable to investigation, than scientific theorizing which, for Fodor, is the paradigm case of a central thought process. Like all other models of human inference, the sketch we offer is tentative: The evidence available so far is compatible with very different approaches. However, as we try to show, the requirement that such a model should help account for inferential communication is both constraining and suggestive.

**2.1. Nondemonstrative inference and strength of assumptions.** *Inference* is the process by which an assumption is accepted as true or probably true on the strength of the truth or probable truth of other assumptions. In *demonstrative inference*, the only form of inference that is well understood, the truth of the premises guarantees the truth of the conclusions. In *nondemonstrative inference*, the truth of the premises merely makes the truth of the conclusions probable. Clearly, the process of inferential comprehension is nondemonstrative: The evidence provided by the communicator never amounts to a proof of her informative intention.

According to what may be called the "logical view" of human nondemonstrative inference, every assumption resulting from such an inference consists of two representations. The first is a representation of a state of affairs: for instance, (2a) below. The second is a representation of the probability or confirmation value of the first representation: for instance, (2b):

- (2) (a) Jane likes caviar.
- (b) The confirmation value of (a) is 0.95.

How are these two representations arrived at? The first, so the story goes, is the output of a nonlogical cognitive process of assumption formation. The second is the output of a process of logical computation which takes as input the assumption to be confirmed on the one hand, and the available evidence on the other.

According to the "functional view," which we put forward, an assumption consists of a single representation, such as (2a). The confidence with which this assumption is held – what we call its *strength* – is a result of its processing history and not of some special computation. The initial strength of an assumption depends on the way it is acquired. For instance, assumptions based on a clear perceptual experience tend to be very strong; assumptions based on the acceptance of somebody's word have a strength commensurate with one's confidence in the speaker; the strength of assumptions arrived at by deduction depends on the strength of the premises from which they were derived. Thereafter, it might be that the strength of an assumption is increased every time it helps in processing some new information and is diminished every time it makes the processing of new information more difficult. According to this view, the strength of an assumption is a functional property, just like, say, its accessibility; it need not be represented in the mind (though it can be). We argue moreover that the strength

of an assumption, unlike its confirmation value, is, in the terms of Carnap (1950), a comparative rather than a quantitative feature: It allows only gross absolute judgments, and finer comparisons of closely related cases.

**2.2. Deduction and its role in nondemonstrative inference.** By its very definition, a nondemonstrative inference cannot *consist* in a deduction. Many authors seem to make the much stronger and unwarranted hypothesis that a nondemonstrative inference cannot *contain* a deduction as one of its subparts. The recovery of implicatures, for example, is a paradigm case of nondemonstrative inference, and it is becoming a commonplace of the pragmatic literature that deduction plays little if any role in this process.<sup>7</sup> We maintain, on the contrary, that the spontaneous and essentially unconscious formation of assumptions by deduction is a key process in nondemonstrative inference. More generally, the ability to perform deductions provides the mind with a uniquely adapted means of extracting more information from information it already possesses, of deriving the maximum cognitive benefit from new information, and of checking the mutual consistency of its assumptions.

To model this unconscious deductive ability, we describe a *deductive device* which takes as input a set of assumptions and systematically deduces all the conclusions it can from them. If this device were equipped with a standard logic, it would derive an infinity of conclusions from any given set of premises; its operations would therefore never come to an end. However most of these conclusions would be of a trivial sort (e.g.,  $(P \text{ and } Q)$ ,  $((P \text{ and } Q) \text{ and } Q)$ , deduced from  $P$  and  $Q$  by the standard rule of *and*-introduction). On the other hand, many deductions which do play a role in ordinary thinking would not be made at all; for example, the deduction from premises (3) and (4a) to the conclusion (5) would not be performed, because it requires a nonstandard rule of concept logic allowing (4b) to be deduced from (4a):

(3) If a relative of Peter's was present, he must have been happy.

(4) (a) Peter's mother was present.

(b) A relative of Peter's was present.

(5) Peter must have been happy.

One gets a more adequate picture of human deductive abilities by assuming that the rules available to the deductive device are not those of a standard logic but are *elimination rules* attached to concepts. We treat concepts as triples of (possibly empty) entries – logical, lexical, and encyclopaedic – filed at a single address. The *logical entry* of a concept consists of deductive rules that apply only to sets of premises in which that concept occurs, yielding only conclusions from which that occurrence has been eliminated. Examples of such elimination rules are the standard *and*-elimination rule, or *Modus ponendo ponens* (eliminating “if . . . then”), and the rules of concept logic which determine deductions from “he ran” to “he moved,” from “the glass is red” to “the glass is coloured,” or from (4a) to (4b). The *encyclopaedic entry* of a concept contains information about the objects, events, or properties that instantiate it. The *lexical entry* contains information about the word or phrase of natural language that expresses the concept. The *address* of a concept, when it

appears in the logical structure of an assumption, gives access to these three types of entry.

We show how the deductive device, drawing on elimination rules attached to concepts, will, from a finite set of premises, automatically deduce a finite set of nontrivial conclusions. We distinguish not only trivial from nontrivial implications, but also analytic from synthetic implications, and discuss their respective role in comprehension. We are particularly concerned with the effect of deductions in which the initial set of assumptions placed in the memory of the deductive device can be partitioned into two subsets, corresponding respectively to some item of new information and to the context in which the new information is processed. Such a deduction may yield conclusions not derivable from either the *new information* or the *context alone*. These we call the *contextual implications* of the new information in the context. A contextual implication is thus a synthesis of old and new information. We see it as a central function of the deductive device to derive, spontaneously, automatically and unconsciously, the contextual implications of any newly presented information in a context of old information.

The information processed by the deductive device, whether new and derived from input systems or old and derived from memory, comes in the form of assumptions with variable strength. We discuss in detail how conclusions inherit their strength from premises. This allows us to characterise three types of *contextual effects* that the processing of new information in a context may bring about: the first, already considered, is the derivation of new assumptions as contextual implications; the second is the strengthening of old assumptions; and the third is the elimination of old assumptions in favour of stronger new assumptions which contradict them. The notion of a contextual effect is essential to a characterisation of relevance.

### 3. Relevance

**3.1. Degrees of relevance: Effect and effort.** It should be clear that we are not trying to define the ordinary and rather fuzzy English word *relevance*. We believe, though, that there is an important psychological property – a property involved in mental processes – which the ordinary notion of *relevance* roughly approximates, and which it is therefore appropriate to call by that name, using it in a technical sense.

As we show in the book, the notion of a contextual effect can be used to state a necessary and sufficient condition for relevance: An assumption is *relevant* in a context if, and only if, it has some contextual effect in that context. This captures the intuition that, to be relevant in a context, an assumption must connect with that context in some way; and it clarifies this intuition by specifying the nature of the connection required. Such a definition, however, is insufficient for at least two reasons: the first is that relevance is a matter of degree and the definition says nothing about how degrees of relevance are determined; the second reason is that it defines relevance as a relation between an assumption and a context, whereas we might want to be able to describe the relevance of any kind of information to any kind of information-processing device,



and more particularly to an individual. At the moment, then, we have simply defined a formal property, leaving its relation to psychological reality undescribed.

Consider first the question of degrees of relevance. What we propose is a kind of cost/benefit analysis. We argue that the contextual effects of an assumption in a given context are only one of two factors to be taken into account. Contextual effects are brought about by mental processes; mental processes, like all biological processes, involve a certain effort. This processing effort is the second factor involved. We then define:

*Relevance:*

*Extent condition 1:* An assumption is relevant in a context to the extent that its contextual effects in that context are large.

*Extent condition 2:* An assumption is relevant in a context to the extent that the effort required to process it in that context is small.

This definition of relevance is comparative rather than quantitative. It makes clear comparisons possible only in some cases: Other things being equal, an assumption with greater contextual effects is more relevant; and, other things being equal, an assumption requiring a smaller processing effort is more relevant. When effect and effort vary in the same direction, comparison may be impossible.

Relevance could be defined not just as a comparative but as a quantitative concept, which might be of some interest to logicians and AI specialists. However, the notion needed by psychologists is the comparative one. It is highly implausible that individuals *compute* the size of cognitive effects and mental efforts. We assume rather that the mind assesses its own efforts and their effects by monitoring physico-chemical changes in the brain. We argue then that effect and effort are nonrepresentational dimensions of mental processes: That is, they exist whether or not they are represented; and when they are represented, it is in the form of intuitive comparative judgments. The same holds for relevance, which is a function of effect and effort.

**3.2. The relevance of a phenomenon to an individual.** In much of the pragmatic literature, relevance is seen as a variable to be assessed in a predetermined context. This is psychologically unrealistic. As we amply illustrate, the context is not given but chosen. Moreover, humans are not in the business of simply assessing the relevance of new information. They try to process information as relevantly as possible; that is, they try to obtain from each new item of information as great a contextual effect as possible for as small as possible a processing effort. For this, they choose a context which will maximise relevance. In verbal communication in particular, relevance is more or less treated as given and context is treated as a variable.

At any moment, an individual has at his disposal a particular set of accessible contexts. There is first an *initial context* consisting of the assumptions used or derived in the last deduction performed. This initial context can be expanded in three directions: by adding to it assumptions used or derived in preceding deductions, by adding to it chunks of information taken from the encyclopaedic entries of concepts already present in the

context or in the assumption being processed, and by adding input information about the perceptual environment. Thus each context except the initial one includes other contexts: The set of accessible contexts is partly ordered by the inclusion relation. This formal relation has a psychological counterpart: Order of inclusion corresponds to order of accessibility.

Treating relevance as a property of propositions or assumptions (as is often done in the pragmatic literature) involves a considerable abstraction: Individuals do not directly pick up an assumption from an utterance, or, more generally, from perceptible phenomena in their environment. Each phenomenon may give rise to a wide range of assumptions or be left unattended. If relevance theory is to explain ostensive-inferential communication, it must explain how attention is directed to a particular phenomenon, and which assumption is likely to be constructed to describe it. For this, we need to define not just the relevance of an assumption, but more generally the relevance of a phenomenon.

Note that the choice, or construction, of an adequate context by expanding the initial context requires some effort, and so does the construction of an assumption about a phenomenon on the basis of the sensory stimulation it provides. To convert our definition of the relevance of an assumption in a context into a definition of the relevance of a phenomenon to an individual, all we have to do, then, is add on the effort side the effort required to arrive at an assumption and a context:

*Relevance of a phenomenon to an individual:*

*Extent condition 1:* A phenomenon is relevant to an individual to the extent that the contextual effects achieved in processing it are large.

*Extent condition 2:* A phenomenon is relevant to an individual to the extent that the effort required to process it is small.

One could leave this definition as it stands and take the relevance of a phenomenon to vary according to how it is actually processed. We propose instead to take the relevance of a phenomenon to an individual to be the relevance achieved when it is optimally processed: that is, when the best possible representation and context are constructed, and by the most economical method. In *Relevance*, we suggest that this proposal, together with the assumption that human cognition is relevance-oriented, yields new insight into the focusing of attention, the choice of a particular representation for a given phenomenon, and the organisation of memory.

**3.3. The principle of relevance.** When paying attention to an ordinary phenomenon, the individual may have hopes of relevance. What makes these hopes reasonable is that humans have a number of heuristics, some of them innate, others developed through experience, aimed at picking out relevant phenomena. Even so, hopes of relevance sometimes turn out to be unjustified; and when they are justified, they are justified to a greater or lesser extent: There can be no general expectation of a steady and satisfactory level of relevance in individual experience.

With an ostensive stimulus, however, the addressee can have not only hopes but also fairly precise expectations of relevance. It is manifest that an act of ostensive

communication cannot succeed unless the addressee pays attention to the ostensive stimulus. It is also manifest that people will pay attention to a phenomenon only if it seems relevant to them. It is manifest, then – mutually manifest in normal conditions – that a communicator must intend to make it manifest to the addressee that the ostensive stimulus is relevant to him. In other words, an act of ostensive communication automatically communicates a *presumption of relevance*. We argue that the presumption of relevance is different on the effect and effort sides. On the effect side, the presumption is that the level of achievable effects is never lower than is needed to make the stimulus worth processing; on the effort side, it is that the level of effort required is never gratuitously higher than is needed to achieve these effects.

The level of relevance thus presumed to exist takes into account the interests of both communicator and audience. Let us call it a level of *optimal relevance*. We can now spell out the presumption of optimal relevance communicated by every act of ostensive communication:

*Presumption of optimal relevance:*

- (a) The set of assumptions *I* that the communicator intends to make manifest to the addressee is relevant enough to make it worth the addressee's while to process the ostensive stimulus.
- (b) The ostensive stimulus is the most relevant one the communicator could have used to communicate *I*.

What we call the *principle of relevance* is the thesis that every act of ostensive communication communicates the presumption of its own optimal relevance. We argue that the principle of relevance explains how the production of an ostensive stimulus can make the communicator's informative intention mutually manifest, thus leading to the fulfilment of the communicative intention. Several inferential steps are involved. In *Relevance*, we discuss how the stimulus can be recognised as ostensive, and how its structure, in the case of both coded and noncoded stimuli, makes accessible a range of hypotheses about the communicator's informative intention. Here we consider only how the principle of relevance provides a sufficient criterion for selecting one of these hypotheses.

Once the ostensive nature of a stimulus is manifest, it is also manifest that the communicator has the informative intention of making manifest to the addressee some set of assumptions *I*. What the principle of relevance does is identify one member of *I*: namely, the presumption of relevance. The presumption of relevance is not just a member of *I*, it is also *about I*. As a result, it can be confirmed or disconfirmed by the contents of *I*. A rational communicator (who genuinely intends to communicate rather than, say, distract an audience) must expect the identification of *I* to confirm the presumption of relevance. To recognise the communicator's informative intention, the addressee must then discover for which set *I* the communicator had reason to expect that *I* would confirm the presumption of relevance. We argue that this is all he has to do.

Let us say that an interpretation is *consistent with the principle of relevance* if and only if a rational communicator might expect it to be optimally relevant to the addressee. Imagine an addressee who tests hypotheses about the contents of *I* in order of accessibility. Suppose he arrives at a hypothesis that is consistent with the

principle of relevance. Should he stop there, or go on and test the next hypothesis on the ground that it too may be consistent with the principle of relevance? Suppose he does go on, and finds another hypothesis that verifies the first part of the presumption of relevance: The putative set *I* is relevant enough. In these circumstances, the second part of the presumption of relevance is almost invariably falsified: The communicator should have used a stimulus that would have saved the addressee the effort of first accessing two hypotheses consistent with the principle of relevance, and then having to choose between them. Thus, the principle of relevance warrants the selection of the first accessible interpretation consistent with the principle. If there is such an interpretation and it is the one intended, communication succeeds. Otherwise it fails.

The principle of relevance does, with much greater explicitness, all the explanatory work of Grice's maxims, and more. There is, however, a radical difference between the principle of relevance and Grice's maxims. Grice's cooperative principle and maxims are norms which communicators and audience must know in order to communicate adequately. Although communicators generally keep to the norms, they may also violate them to achieve particular effects; and the audience uses its (presumably learned) knowledge of the norms in interpreting communicative behaviour.

The principle of relevance, by contrast, is a generalisation about ostensive-inferential communication. Communicators and audience need no more know the principle of relevance to communicate than they need to know the principles of genetics to reproduce. It is not the general principle but the fact that a particular presumption of relevance has been communicated, by and about a particular act of communication, that the audience uses in inferential comprehension. Communicators do not "follow" the principle of relevance; and they could not violate it even if they wanted to. The principle of relevance applies without exception: Every act of ostensive communication communicates a presumption of relevance. Note, though, that the presumption of relevance carried by a particular act of communication does not have to be true or accepted as true: The communicator may fail to be relevant. It is enough that the presumption of relevance should be communicated – and it always is – to fulfil its most important role: determining the interpretation of the ostensive stimulus.

#### 4. Aspects of verbal communication

Verbal communication, we argue, involves two types of communication process: one based on coding and decoding, the other on ostension and inference. The coded communication is of course linguistic: A linguistic stimulus triggers an automatic process of decoding. The semantic representations recovered by decoding are *logical forms* which, like the logician's open sentences, but in more ways, fall short of determining a single proposition. These logical forms, we claim, never surface to consciousness. Instead, they act as assumption schemas which can be inferentially completed into fully *propositional forms*, each determining a single proposition and serving as a tentative identification of the intended ex-



PLICIT content of the utterance. This explicit content alone has contextual effects and is therefore worthy of conscious attention.

The coded communication process, then, serves as a source of hypotheses and evidence for the second communication process, the inferential one. If comprehension is defined as a process of identifying the speaker's informative intention, linguistic decoding is better seen not as part of comprehension proper, but rather as providing the main input to the comprehension process. Unlike most pragmaticians, who see the inferential tier of comprehension as governed by a variety of specialised rules constituting a kind of pragmatic "module," we argue that it involves only the application of nonspecialised inference rules – rules that apply as well to all conceptually represented information. We see pragmatics as the study, not of a distinct mental device, but of the interaction between a linguistic input module and central inferential abilities.<sup>8</sup>

**4.1. The identification of explicit content.** The first task in inferential comprehension is to complete the logical form recovered by decoding and identify the explicit content of the utterance.<sup>9</sup> This in turn involves three subtasks: disambiguation (when, as is usual, the decoding yields a choice of logical forms); identification of the referents of referring expressions; and enrichment of the schema selected – a subtask less often considered, of which we analyse several examples.

These subtasks could in principle yield a variety of outcomes. By what criterion does the hearer recognise the right explicit content, that is, the one he was intended to choose? Although there is a considerable literature on disambiguation and reference assignment, this question has not been seriously addressed. The aims of psycholinguists lie elsewhere: They want to describe not the criteria used in disambiguation, but the procedure by which it is achieved. The only criterion generally considered, and apparently confirmed by much experimental evidence, is one of economy, something like a principle of least effort.

Could the answer simply be that the right explicit content is the one obtained by going through some effort-saving procedure? The existence of so-called garden-path utterances (e.g., "I saw that gasoline can explode. And a brand new can it was too") strongly suggests that whatever regular procedures are available for disambiguation, reference assignment, and enrichment yield at best a tentative identification, one that will be rejected if it turns out not to meet some as yet unspecified criterion. We show, with examples, that the correct criterion is consistency with the principle of relevance. This answer does not eliminate considerations of effort; on the contrary, it integrates them by suggesting that the least effort-consuming, and therefore potentially the most relevant interpretation, should be considered first (although it should be abandoned if it fails to yield the expected effect).

**4.2. The identification of implicatures.** We introduce a distinction between two kinds of implicatures: *implicated premises* and *implicated conclusions*. Implicated conclusions are deduced from the explicit content of an utterance and its context. What makes it possible to identify

such conclusions as implicatures is that the speaker must have expected the hearer to derive them, or some of them, given that she intended her utterance to be manifestly relevant to the hearer. Implicated premises are added to the context by the hearer, who either retrieves them from memory or constructs them ad hoc. What makes it possible to identify such premises as implicatures is that the speaker must have expected the hearer to supply them, or some of them, in order to be able to deduce the implicated conclusions and thereby arrive at an interpretation consistent with the principle of relevance.

To illustrate, consider dialogue (6):

- (6) (a) *Peter*: Would you drive a Mercedes?  
(b) *Mary*: I wouldn't drive ANY expensive car.

The explicit content of Mary's reply does not directly answer Peter's question. However, processed in a context containing (7), (6b) yields the contextual implication (8):

- (7) A Mercedes is an expensive car.  
(8) Mary would not drive a Mercedes.

We have a situation, then, in which Mary, in producing (6b), has not directly and explicitly answered Peter's question, but has made manifest a contextually implied answer, that is, (8). Given that she could not expect her utterance to be relevant unless it made manifest such an answer, this implied answer is manifestly intentional: It is an implicated conclusion of her utterance. Since it is manifest that Peter would not have deduced this conclusion without adding (7) to the context, (7) is an implicated premise of Mary's utterance.

Implicatures (7) and (8) have two properties which many pragmaticians think of as shared by all implicatures. In the first place, they are fully determinate. Mary expects Peter to supply not merely something *like* premise (7) and conclusion (8), but a premise and conclusion with just this logical content. Second, Mary guarantees their truth. Suppose that before (6b) was produced, Peter had not known that Mercedes cars were expensive; then (6b) would give him as much reason to think they are as if Mary had explicitly asserted it. There has been a tendency in modern pragmatics to treat *all* implicatures as fully determinate assumptions for which the speaker is just as responsible as if she had asserted them directly. According to this approach, utterance comprehension consists in the recovery of an enumerable set of assumptions, some explicitly expressed, others implicitly conveyed, but all individually intended by the speaker. We argue that this is a mistake, or, at best, a counterproductive idealization.

Consider, for instance, the exchange in (9):

- (9) *Peter*: What do you intend to do today?  
*Mary*: I have a terrible headache.

What does Mary implicate? That she will not do anything? That she will do as little as possible? That she will do as much as she can? That she does not yet know what she will do? There is no precise assumption, apart from the one explicitly expressed, which she can be said to intend Peter to share. Yet there is more to her utterance than its explicit content; she manifestly intends Peter to draw some conclusions from what she said, and not just any conclusions. Quite ordinary cases such as (9) are

never discussed in the pragmatic literature. Moreover, even in cases such as (6)–(8), the implicit import of the utterance is usually not exhausted by clear implicatures such as (7) and (8). Mary's reply (6b) suggests, but only in a vaguer way, that she would not drive, say, a Jaguar, that she finds driving an expensive car objectionable, and so on. What pragmatics needs – and relevance theory provides – is a precise account of these vaguer effects.

In our framework, the greater the mutual manifestness of the informative intention to make manifest some particular assumption, the more strongly this assumption is communicated. Using this approach, the indeterminacy of implicatures presents no particular formal problem. An utterance that forces the hearer to supply a very specific premise or conclusion to arrive at an interpretation consistent with the principle of relevance has a very strong implicature. An utterance that can be given an interpretation consistent with the principle of relevance on the basis of different – though of course related – sets of premises and conclusions has a wide range of weak implicatures. Clearly, the weaker the implicatures, the less confidence the hearer can have that the particular premises or conclusions he supplies closely reflect the speaker's thoughts, and this is where the indeterminacy lies. However, people may entertain different thoughts and come to have different beliefs on the basis of the same cognitive environment. The aim of communication in general, we claim, is to increase the mutuality of cognitive environments and thereby the similarity of thoughts, rather than to guarantee a (generally unreachable) strict duplication of thoughts.

#### 4.3. Explicit content and style: Presuppositional effects.

It might seem that two utterances with the same linguistically determined truth-conditions must have identical implicatures. We argue that this is not so: They may differ in the processing effort they require, which, given the principle of relevance, can lead to different effects. This, we claim, is the key to an explanatory theory of style. In *Relevance*, we discuss two classes of stylistic effects: presuppositional effects at some length and poetic effects more briefly. Here we will merely indicate the relevance of relevance theory to these two aspects of style.

Presuppositional effects result from the fact that a sentence is not decoded as a single symbol, in one go, but in steps, as a structured string of constituents (some of which may be highlighted by stress – an important aspect left out of this précis). Each constituent provides some information by allowing analytic implications to be deduced. Among these, we distinguish *foreground implications*, which contribute to relevance by having contextual effects, and *background implications*, which contribute to relevance by saving effort. Background implications save effort in particular by making more accessible the context in which foreground implications will produce contextual effects. In general (and we show that exceptions are predicted by relevance theory), an optimally relevant utterance will have its effort-saving background implications made available by initial constituents, and its effect-carrying foreground implications made available by its final constituents: Thus the construction of the context will be well under way, or even over, when the

last word is uttered, and effect can be achieved at the smallest processing cost.

Processing an utterance as economically as possible, the hearer normally treats the implications made available by the initial constituents as background, the implications made available by the last constituents as foreground, and expects the foreground implications to carry their effects in the context made accessible by the background implications. Thus two utterances with the same truth-conditions but different word order lead to the construction of different contexts and the search for different effects. Compare for instance:

- (10) Leo sold Peter a painting.  
 (11) Peter bought a painting from Leo.

In (10) the hearer's expectation is that the utterance will be relevant in a context of information about Leo. If it were mutually manifest, for instance, that Leo desperately needed money, a key implicature would be that he has just made some. In (11), the hearer's expectation is that the utterance will be relevant in a context of information about Peter; for instance, if it were mutually manifest that Peter did not care for Leo's painting but knew he needed money, (11) would implicate that Peter behaved generously. In other words, even though (10) and (11) have the same truth-conditions and the hearer could in principle draw the same contextual implications from either, not all of the implications are implicated, or implicated to the same degree, by both utterances. This is because the two utterances organise the hearer's efforts differently.

We argue in *Relevance* that all the stylistic effects discussed in the literature in terms of presupposition and focus, presupposition and assertion, topic and comment, given and new, theme and rheme, and so on,<sup>10</sup> can be explained with greater generality, simplicity, and predictive power in terms of background and foreground. Unlike these other distinctions, which purport to describe linguistic or pragmatic properties registered by a competent speaker, the foreground/background distinction is not something that speakers need to have built into either their grammar or their inferential abilities. Given that utterances have constituent structure, internal order, and focal stress and are processed over time, backgrounding and foregrounding arise as automatic effects of the hearer's tendency to maximise relevance and the speaker's exploitation of that tendency.

**4.4 Implicatures and style: Poetic effects.** Style arises, we maintain, in the pursuit of relevance. In allocating the information she wants to communicate between the explicit content and the implicatures of her utterance, in relying on stronger or weaker implicatures, the speaker makes manifest her assessment of the tenor and quality of mutual understanding between her audience and herself. She thereby gives her utterance its particular style.

Here we will briefly illustrate this point with an example of the simplest of all the classical figures of speech: epizeuxis or repetition. Compare the interpretation of (12) and (13):

- (12) My childhood days are gone.  
 (13) My childhood days are gone, gone.

Both have the same truth-conditions and therefore potentially the same contextual implications. What (13) has is more *implicatures* than (12): that is, more contextual assumptions and implications that receive some degree of backing from the speaker. The repetition of “gone” causes some extra processing effort. Given the principle of relevance, this extra effort should be justified by some extra effect. Having thought of all the implicatures that the speaker could reasonably have expected him to derive from the first occurrence of “gone,” the hearer may assume that there is a whole range of still further premises and conclusions which the speaker wants to implicate. For this, he must expand the context. Thus (13) may encourage the hearer to compare the speaker’s childhood and her present condition, to assume that she herself is reminiscing and making a similar comparison, and to imagine the feelings this may evoke in her. What the repetition produces, then, is many very weak implicatures.

We suggest that the peculiar effect of an utterance that achieves most of its relevance through a wide array of weak implicatures is properly called a *poetic effect*. How do poetic effects affect the mutual cognitive environment of speaker and hearer? They do not add entirely new assumptions which are strongly manifest in this environment. Instead, they marginally increase the manifestness of a great many weakly manifest assumptions. In other words, poetic effects create common impressions rather than common knowledge. Utterances with poetic effects can be used precisely to create this sense of apparently affective rather than cognitive mutuality. What we are suggesting is that, if you look at these affective effects through the microscope of relevance theory, you see a wide array of minute cognitive effects.

**4.5. Descriptive and interpretive dimensions of language use.** There is a considerable literature on illocutionary force and speech acts, and an even more considerable one on tropes. There is very little overlap between the two, as if it went without saying that these are two essentially different aspects of language use. In both cases, the literature is centrally concerned with problems of classification and offers little in the way of explanation. We propose a new, more integrated and more explanatory approach, based on a fundamental distinction between interpretation and description.<sup>11</sup>

The relationship between a representation and the object it represents can be of two kinds: It can be based on resemblance or on truth. Any object in the world can, under appropriate conditions, be used as a representation of some other object that it resembles. You ask me what is the shape of Brazil, and by way of reply, I point to an appropriately shaped cloud in the sky. Resemblance raises well-known philosophical and psychological problems: How are the pertinent features of the representation identified? How are degrees of resemblance assessed? That is, how can the representation inform one about the object it resembles? At least when resemblance is used ostensively, relevance theory provides the key to an answer: Consider hypotheses in their order of accessibility (salient features first, etc.) and select the first hypothesis that the producer of the representation may have thought would be relevant enough.

An object with a propositional content – an utterance, for example – can be used to represent in two quite different ways. It can represent some state of affairs by virtue of being true of that state of affairs; in this case we will say that the representation is a *description*, or that it is used *descriptively*. Or, like any object, it can represent something it resembles, and in particular some other representation with a similar propositional content; in this case we will say that the first representation is an *interpretation* of the second one, or that it is used *interpretively*. Two representations resemble one another interpretively when they share analytic and contextual implications.<sup>12</sup>

The only generally acknowledged interpretive use of utterances is in the reporting of speech or thought, as in quotations and summaries. However, there are others. Utterances can be used interpretively to represent utterance-types, or thoughts worth considering for their intrinsic properties, rather than because they can be attributed to Peter, Mary, or public opinion. We argue that there is an even more essential interpretive use of utterances: On a more fundamental level, *every* utterance is used interpretively to represent a thought of the speaker’s. One of the assumptions a speaker intends to make manifest is that she is entertaining a thought with some particular attitude: It is on this ground that the hearer may be led to entertain a similar thought with a similar attitude. You may well tell me that you will come tomorrow, but you will not make me believe it unless you first make me believe that you believe it too. This much is hardly controversial. In our terms, it means that an utterance is, or purports to be, in the first instance, an interpretation of a thought of the speaker.

Actually, an even stronger claim is generally made. Most pragmaticians and philosophers of language take for granted that there is a convention, principle, or presumption to the effect that the propositional content of the utterance must be a literal expression – that is, a strictly faithful interpretation – of a thought of the speaker’s.<sup>13</sup> We argue that this claim is too strong. How close the interpretation is, and in particular when it is literal, can be inferentially determined by the hearer.

What does the thought interpretively represented by an utterance itself represent, and how? A mental representation, like any representation, can be used descriptively or interpretively. When it is used descriptively, it can be a description of an actual state of affairs, or it can be a description of a desirable state of affairs. When it is used interpretively, it can be an interpretation of an attributed thought or utterance, or it can be an interpretation of a relevant, hence desirable thought. There may be other possibilities, and one might consider what the thoughts interpreted by thoughts might represent in their turn and how; but let us leave it at that, and use the following diagram to show the representations and relationships considered so far.

Any utterance involves at least two levels of representation: It interpretively represents a thought of the speaker’s, which itself descriptively represents some state of affairs, or interpretively represents some further representation. All the basic relationships involved in tropes and illocutionary forces are represented in Figure 1: Metaphor involves an interpretive relation between the



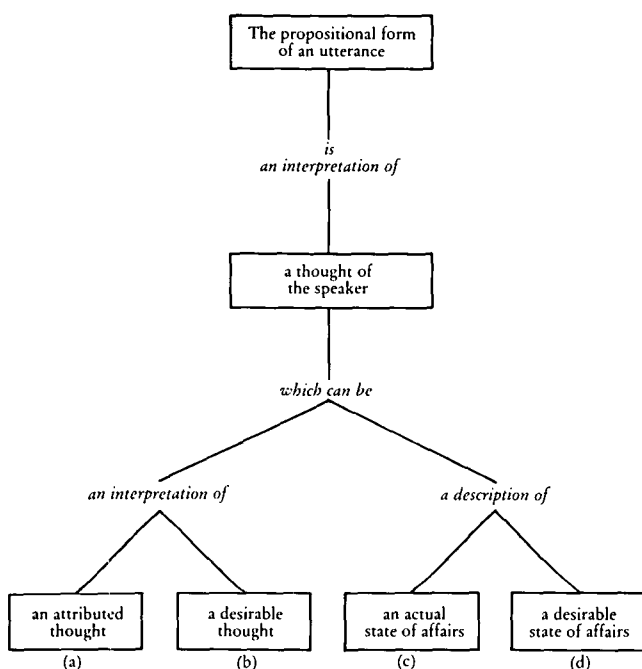


Figure 1. Descriptive and interpretive dimensions of language use.

propositional form of an utterance and the thought it represents; irony involves an interpretive relation between the speaker's thought and attributed thoughts or utterances; assertion involves a descriptive relation between the speaker's thought and a state of affairs in the world; requesting or advising involves a descriptive relation between the speaker's thought and a desirable state of affairs; interrogatives and exclamatives involve an interpretive relation between the speaker's thought and a relevant – that is, desirable – thought. Our book develops all of these ideas; here we shall briefly consider only the potentially more controversial interpretive cases: those of metaphor, irony, interrogatives, and exclamatives.

**4.6. Literalness, looseness, and metaphor.** An utterance, we claim, is an interpretation of one of the speaker's thoughts, that is, both share a number of implications. This means that, in order to communicate a set of assumptions *I* which are the essential implications of some thought of hers, the speaker must produce an utterance whose explicit content logically or contextually implies *I*. An utterance that implies exactly *I* and nothing else may not be available, or it may be available but not economical. However, the speaker may freely use an utterance that implies many other assumptions she does not want to endorse, as long as the hearer has some way of recognising the intended implications which are members of *I*. In our framework, all the hearer has to do is start computing, in order of accessibility, those implications which are, or which the speaker would consider to be, relevant to him, and continue to add them to the overall interpretation of the utterance until it is, or might have seemed to the speaker, relevant enough to be consistent with the principle of relevance. At this point, the sorting will have been accomplished as a byproduct of the search for relevance: It will require no specific effort of its own.

The comprehension of every utterance involves such a

process of identifying relevant implications. When the proposition expressed is itself among the implications on which optimal relevance depends, the result is a literal interpretation. According to this view, literalness is simply maximal resemblance, and enjoys no privileged status. In general, though, some looseness of expression is to be expected. For example, the speaker may spare the hearer some processing effort, and thereby optimize relevance, by saying "It's half-past five" rather than "It's twenty-eight minutes past five" (even though she knows that the latter is true) when the relevance-producing effects of her utterance do not depend on a strictly literal reading. Similarly, if the speaker says "I'm exhausted," no one will stop to wonder whether exhausted is exactly what she is, as long as there is an acceptable range of implications which makes her utterance relevant enough. This is again true if the speaker says, in the same circumstances, "I'm dead" – a clear metaphor this time.

In *Relevance*<sup>14</sup> we show that there is no discontinuity between literal uses, loose uses, and metaphors. According to this approach, metaphor (like a variety of related tropes such as hyperbole, metonymy, and synecdoche) requires no special interpretive abilities or procedures: The fact that some utterances are interpreted metaphorically, just as others are interpreted literally, results from the same standard process of comprehension.

**4.7. Echoic utterances and irony.** We argue that irony and a variety of related tropes such as meiosis and litotes fall together with a range of cases not normally regarded as figurative at all. What unites these cases is the fact that the thought of the speaker, which is interpreted by the utterance, is itself an interpretation. What it interprets is a thought of someone (or some group) other than the speaker – or of the speaker in the past. That is, these utterances are second-degree interpretations of someone else's thought, as illustrated by path (a) in Figure 1 above.

How do interpretations of someone else's thought achieve relevance? In the best-known case, that of "reported speech," they achieve relevance by informing the hearer of the fact that so-and-so has said something or thinks something. In other cases, these interpretations achieve relevance by informing the hearer of the fact that the speaker has in mind what some individual or individuals say or think and has a certain attitude to it. When interpretations achieve relevance in this way, we say that they are *echoic*.

By representing a thought that is not her own in a manifestly sceptical, amused, surprised, triumphant, approving, or reproving way, the speaker can express her own attitude to it. There is no limit to the attitudes that a speaker can express to an opinion echoed. In particular, she may indicate her agreement or disagreement. Compare (14) and (15):

- (14) (a) *He*: It's a lovely day for a picnic.  
(They go for a picnic and the sun shines.)  
(b) *She* (happily): A lovely day for a picnic, indeed!  
(15) (a) *He*: It's a lovely day for a picnic.  
(They go for a picnic and it rains.)  
(b) *She* (sarcastically): A lovely day for a picnic, indeed!

In both (14b) and (15b) there is an echoic allusion to be picked up. In the circumstances described, it is clear that the speaker of (14b) endorses the opinion echoed, whereas the speaker of (15b) rejects it with scorn. These

utterances are interpreted on exactly similar patterns, the only difference being in the attitudes they express: (14b) has not been thought by rhetoricians to be worthy of special attention; (15b) is, of course, a case of verbal irony.

An ironical attitude is of the disapproving kind. From the display of such an attitude, the hearer can infer, if it was not already obvious to him, that the speaker believes the opposite of the opinion echoed: Thus the speaker of (15b) manifestly believes that it is *not* a lovely day for a picnic. However, against classical approaches to irony, we argue that this is not the “figurative meaning” of the ironical utterance. At most, it is one of its implicatures: More relevant implicatures might be that her companion’s judgment has been unsound, that they should never have set out, that it was his fault that their day has been ruined, and so on. The recovery of these implicatures depends, first, on a recognition of the utterance as echoic, second, on an identification of the source of the opinion echoed, and third, on a recognition that the speaker’s attitude to the opinion echoed is one of disapproval. We argue that these are the key factors in the interpretation of all ironical utterances.<sup>15</sup>

We are arguing on the one hand that metaphors and ironies are not essentially different from other types of utterances but that, on the other hand, they are not essentially similar to one another. Metaphor plays on the relationship between the propositional form of an utterance and that of the speaker’s thought; irony plays on the relationship between the speaker’s thought and a thought it interprets. This suggests that the classical notion of a trope, which covers metaphor and irony and distinguishes both from “nonfigurative” utterances, should be abandoned altogether: It groups together phenomena which are not closely related and fails to group together phenomena which are.

**4.8. Speech acts: Interrogatives and exclamatives.** Our book questions some of the basic assumptions of current speech-act theory, and sketches an alternative approach which puts a much greater load on inference than on decoding in the identification of illocutionary force. Given the principle of relevance, we argue that illocutionary-force indicators such as declarative or imperative mood or interrogative word order have to make manifest only a rather abstract property of the speaker’s informative intention: the direction in which the relevance of the utterance is to be sought. Here, we will take the case of interrogatives and exclamatives as an illustration of this general approach.

Speech-act theorists tend to analyse interrogative utterances as a special subtype of directive speech act: specifically, as requests for information (see Bach & Harnish 1979, p. 48; Searle 1969, p. 69). As a result, rhetorical questions such as (16), expository questions such as (17), and self-addressed questions such as (18), all require ad hoc separate treatment:

- (16) When did you say you were going to give up smoking?
- (17) What are the main objections to this approach? First . . .
- (18) Why do we have to die?

We argue that what the interrogative word order signals is a much more abstract property than that of being a request for information. It signals that the utterance represents (not descriptively but interpretively, of course) an assumption which would be relevant if true. In

other words, a question is an interpretation of a desirable thought – the piece of knowledge which would answer it; it follows path (b) on Figure 1. In the case of a yes–no interrogative, the propositional form of the utterance is like a quotation of the relevant-if-true assumption. In the case of a wh-question (who-which-what-why and so on), the logical form of the utterance is not fully propositional; it is an incomplete interpretive representation of the relevant-if-true assumption.

On this account, this is all the interrogative word order signals. Decisions about who would find the assumption represented relevant, and about whether the speaker expects an answer, are left to the hearer to infer on the basis of the principle of relevance. For example, rhetorical questions such as (16) are often reminders, designed to prompt the retrieval of an assumption the speaker regards as relevant to the hearer. Expository questions such as (17) are analysable as questions whose answers the speaker not only regards as relevant to the hearer but is about to provide herself. Regular requests for information, by contrast, are analysable as questions whose answers the speaker regards as relevant to her and, moreover, expects the hearer to supply. In pure speculations such as (18), the suggestion is that the answer would be relevant to the speaker, or to both speaker and hearer, but there is no expectation that anyone will be in a position to supply it. There is thus no need to analyse all questions as requests for information, no need to set up special speech-act categories to handle offers of information, rhetorical questions, expository questions, and so on.

One advantage of this approach is that it suggests a way of explaining the striking syntactic parallels between interrogative and exclamative sentences such as (19) and (20):

- (19) How clever is Jane?
- (20) How clever Jane is!

In traditional speech-act terms, since interrogatives are requests for information and exclamatives are emphatic assertions, it is hard to account for the consistent cross-linguistic parallelisms (see Grimshaw 1979, Sadock & Zwicky 1985) between these two utterance types. On our approach, exclamatives, like interrogatives and unlike declaratives, are specialised for interpretive rather than descriptive use, and also follow path (b) of Figure 1. Whereas a speaker who asks a question such as (19) indicates that some true completion of the incomplete thought represented by her utterance is relevant, a speaker who produces an exclamation such as (20) indicates that some relevant completion of the incomplete thought represented by her utterance is true. In other words, the speaker of (20) indicates that Jane is high enough on the scale of cleverness for this to be worth drawing attention to. Thus, the intuition that exclamatives are like emphatic assertions and the striking parallels between exclamative and interrogative form are simultaneously explained.

## 5. Concluding remark

We are well aware that the view developed in *Relevance* and summarised here is very speculative and, as it stands, too general to determine directly either specific experi-

mental tests or computer simulations. In assessing a new approach to human communication, however, the following questions should be kept in mind. How does it compare with other current approaches in terms of explicitness, plausibility, generality, and explanatory power? Does it throw new light both on the very rich and diverse data available to all of us as individuals involved in communication and on the narrower but more reliable data gathered by scholars? Does it suggest new empirical research? Is it relevant to more than one of the many disciplines involved in the study of human communication – linguistics, pragmatics, philosophy, cognitive psychology, artificial intelligence, social psychology, literary studies, anthropology, and sociology – and could it foster fruitful interactions among them?

## NOTES

1. Correspondence to Dan Sperber should be addressed to 2, Square de Port-Royal, 75013 Paris, France.
2. See, for example, Bach and Harnish (1979); Clark (1977); Clark and Carlson (1981); Green and Morgan (1981); Leech (1983); Lewis (1979); Sag (1981).
3. Notably Gazdar (1979). For recent surveys of the pragmatics literature, see Brown and Yule (1983); Levinson (1983).
4. For a general discussion of the mutual knowledge issue, see Neil Smith (1982), in particular contributions by Clark and Carlson, Johnson-Laird, and Sperber and Wilson.
5. See Armstrong (1971); Bach and Harnish (1979); Bennett (1976); Blackburn (1984); Davidson (1984); Davies (1981); Grice (1957; 1968; 1969; 1982); Harman (1968); Lewis (1969); Loar (1976; 1981); McDowell (1980); Patton and Stampe (1969); Récanati (1979; forthcoming); Schiffer (1972); Searle (1969; 1983); Strawson (1964; 1969; 1971); Wright (1975); Yu (1979); Ziff (1967).
6. For ease of exposition, we will talk of a female communicator and a male addressee.
7. See Leech (1983, pp. 30–31); Levinson (1983, pp. 115–16); Bach and Harnish (1979, pp. 92–93); Brown and Yule (1983, p. 33); de Beaugrande and Dressler (1981, pp. 93–94).
8. On the issue of modularity, see also Wilson and Sperber (1986b).
9. In *Relevance*, we do not talk of “explicit content”; but we distinguish the “propositional form” of an utterance and its “explicatures.” We also propose a redefinition of the explicit/implicit contrast, and provide both a classificatory and a comparative criterion of explicitness. For reasons of space, we leave these issues out of this précis.
10. See Brown and Yule (1983, chaps. 3–5); Chafe (1976); Clark and Haviland (1977); Givón (1979); Halliday (1967–68); Jackendoff (1972); Lyons (1977, chap. 12.7); Oh and Dinneen (1979); Prince (1981); Reinhart (1981); Rochemont (1988); Taglicht (1984, chaps. 1–3).
11. A distinction relevant also to the philosophy of the social sciences, as argued in Sperber (1985).
12. See Sperber and Wilson (1986b) for further discussion.
13. See Bach and Harnish (1979); Lewis (1975); Searle (1969).
14. See also Sperber and Wilson (1986b).
15. For further discussion see Sperber and Wilson (1981); Sperber (1984). For an experimental approach, see Jorgensen, Miller, and Sperber (1984).

## Open Peer Commentary

*Commentaries submitted by the qualified professional readership of this journal will be considered for publication in a later issue as*

*Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged. All page references are to Sperber and Wilson's Relevance unless otherwise indicated.*

## Comparisons with Grice

Jonathan E. Adler

*Department of Philosophy, Brooklyn College, City University of New York, Brooklyn, N.Y. 11210*

It is convenient for me to raise questions for Sperber & Wilson (S&W) by pressing comparisons with Grice's (1975b) original account. Their explication of *relevance* in terms of the extent of contextual effects seems to be a simplification of the maxims under three of Grice's Categories: Quality, Quantity, and Relation. (See, for example, S&W's illustrations and discussion, pp. 120–22). The other pillar of their theory – effort or processing costs – renders explicit expectations that fall under the category of Manner, as well as the basic presumption of obedience to the Cooperative Principle (CP).

In explicitly distinguishing between truthfulness, informativeness, and relevance, Grice allows for priorities among these qualities. Let me assume, following a number of philosophers of language, that truthfulness should have the highest priority. If the assumption is correct, how will S&W capture it, given that they do not treat truth separately?

An illustration of this priority involves a speaker, S, who in the appropriate context responds to someone's question “Do you believe that the Mets will win the Series?” by simply saying no. There is a common tendency to construe such assertions as implying the “opposite”: S disbelieves that the Mets will win or S believes that the Mets will not win. Whereas what S has strictly said carries no such implication: He denies holding a belief; he does not assert that he believes anything else. (His denial has large scope.) Giving a priority to truthfulness or holding the (consonant) presumption of literalness, which S&W reject (p. 230), supports S's denial of responsibility for the listener's drawing the stronger conclusion. However, the stronger conclusion is more relevant in S&W's sense because it has more contextual effects and, given how readily the conclusion is drawn, it appears to be even less of a processing burden than the stricter reading.

In their discussion of responsibility (chap. 4, sect. 4), S&W rightly note that questions of responsibility are matters of degree. Should the speaker have given his answer with the purpose of deliberately misleading the audience, knowing its tendency to overinfer, we might be less sympathetic. However, if his reply was an indirect consequence of S's wanting to speak the precise, appropriate, simplest truth, then, even if S understood this tendency, I do not see that he can be faulted for the implications that others draw, however common. Perhaps this is too legalistic: I do not want to generalize from this case except to claim that, as we move from conversation to public speaking (hearing) to writing (reading), speakers take less responsibility for implications drawn beyond what is literally said. S&W's explicit introduction of processing costs nicely explains this result, since each step in the chain represents diminished demands of cognitive economy. I am puzzled about why they do not appeal to such differences in communicative settings in their discussions of figurative language.

Irony is a case in which speakers clearly are exploiting, not just involving, the hearer's tendency to vacillate between opposites. On p. 240, S&W criticize Grice and others for leaving it unclear why an ironical utterance should be used in the first place. If someone utters to the driver “There's something coming” when no car is coming, the driver slams on the brakes rather than take the utterance as ironic for “There's nothing coming.” What was said is not ironic, S&W insist, but irrational.



However, on Grice's view too, the foul-up is the speaker's fault. The falsity of the utterance cannot be exploited for ironic purposes because it just is not mutually obvious that it is false, especially given the special demands for caution and rapid judgment by the hearer-driver. The violation of the maxim is not conspicuous.

S&W might wonder, as they do in a related context (pp. 200–201), why, even when a maxim is flaunted, the speaker implicates the opposite. The obvious, perhaps facile answer is: Given mutual knowledge of the practice of ironical usage, and presumed obedience to the CP, assuming that the opposite is what the speaker intended is the best explanation of why the speaker said what he did.

In claiming that the order of accessing hypotheses matters to their selection for interpretation, S&W's relevance theory seems to have a determinateness that Grice's theory lacks. At best, Grice has the idea that a speaker obeying the CP will try to formulate his contribution so that the proper hypothesis to interpret it is the easiest; among all appropriate and available ones, that an audience is expected to have. S&W declare that it should be the first hypothesis consistent with relevance.

S&W contrast their aim of optimization with Grice's of maximization: "It is possible to be optimally relevant without being 'as informative as is required' by the current purposes of the exchange" (p. 162). But just as we have to understand the CP as implicitly requiring that speakers make the least processing demands on hearers, only *ceteris paribus*, so too there must be a *ceteris paribus* clause attached to each of the maxims. This is clear, for example, when Grice discusses a "clash" of maxims (Grice 1975b, pp. 69–71) in which the solution involves optimizing maxims under Quantity and Quality.

Contrasting optimizing with maximizing rules, maxims, or directives typically encourages an illusion of determinateness. Thus, Simon's (1979) idea that real information-processing systems must be "satisficers," not maximizers, has been viewed as a hypothesis that is highly specific, an empirical conjecture, and a (nonrational) rationalization for predictable weaknesses in human reasoning. Both the satisficing model and S&W's relevance theory insist that we select the *first* hypothesis that offers a solution to the problem at hand. This appears highly specific or determinate, as well as empirically bold, because we contrast it with the range of alternatives specifying either that we should choose the second, third, and so forth, hypothesis or that we should examine all the hypotheses to choose the best or "maximizing" solution. For any particular, narrowly circumscribed problem, this contrast can make sense. But it makes no sense when we are speaking of the rational adoption of rules, maxims, or directives to be widely applied to a large range of problems under the weakest of finiteness constraints. The proper contrast is between any such satisficing rule and one demanding that we continue to search for solutions even if there is nothing worthwhile that the extra effort promises. This last alternative we can reject a priori relative to the simplest demands of rationality, namely, that we not continue our effortful search once we have reached our goal. The specificity of choosing the first hypothesis derives from the vague and indeterminate demand that efforts should be undertaken only if they are worthwhile: The first hypothesis is just that hypothesis, whatever it happens to be, beyond which we are not rationally justified in going. For if we were so rationally justified, it would no longer be the first appropriate hypothesis.

## Relevant questions

Kent Bach<sup>a</sup> and Robert M. Harnish<sup>b</sup>

<sup>a</sup>Department of Philosophy, San Francisco State University, San Francisco, Calif. 94132 and <sup>b</sup>Department of Philosophy and Department of Linguistics, Program in Cognitive Science, University of Arizona, Tucson, Ariz. 85721

Sperber & Wilson (S&W) find fault with "Gricean" theories of communication and propose a theory of their own. Both their criticisms and their proposals are interesting and provocative, but not entirely clear. In the space available, we can only raise some questions about their theory and make some brief comments on their criticisms.

### The principle of relevance

1. How do speakers and hearers follow the principle of relevance (PR)? According to S&W the PR applies to *every* act of ostensive communication (e.g., p. 158), but the definition of relevance "makes clear comparisons possible only in some cases" (p. 125). Also, since measuring effort is required for determining comparative relevance, how is effort to be measured? All we are told is that "the mind assesses its own efforts and their effects by monitoring these [physico-chemical] changes" (p. 130). If that is possible somehow, then the hearer can assess his own effort. But how does a communicator anticipate the effort expended by the hearer? S&W assert, for example, that the "communicator should have used a stimulus which would have saved the addressee the effort of first accessing two hypotheses consistent with the principle of relevance, and then having to choose between them" (p. 168); but they do not explain how he is to accomplish this.

2. Where do candidate [I]s (sets of assumptions made manifest by informative intentions) come from, beyond those deducible from the meaning of the utterance? S&W say that "the plausibility of some hypotheses may already be manifest in the environment" (p. 167), but just how this occurs and how manifest hypotheses actually come to mind is not explained. S&W accuse us Griceans of "whistling in the dark" (p. 20) on such questions, but they themselves do likewise, as they eventually acknowledge (p. 170).

3. According to the PR, every act of ostensive communication communicates a presumption of its own optimal relevance, and according to that presumption the ostensive stimulus is the most relevant one the communicator could have used to communicate [I]. There is no requirement on the relevance of *what* is communicated, however. Rather, S&W argue that "it is the first interpretation to occur to the addressee that is the one the communicator intended to convey" (p. 169). This raises a difficult question for S&W: How are (nonstandardized) indirect speech acts possible? How can a speaker communicate one thing by way of communicating another if the addressee stops inferring after the first thing occurs to him? More generally, how do S&W explain how the addressee can modify, supplement, or replace the first assumption that comes to mind, even if it is consistent with the PR?

### Ostensive communication

1. Every act of ostensive communication is supposed to communicate a presumption of its own optimal relevance. But how does one identify a stimulus as an instance of ostensive communication in the first place, so that its optimal relevance can be presumed?

2. Does ostensive communication require a communicative intention? First S&W say that it does, when they define communicative intentions (pp. 60–61), and then, when they define ostensive-inferential communication (p. 63), they say that it does not.

3. Moreover, does ostensive communication require uptake on the part of the hearer? S&W's definition does not require uptake, since it specifies only that the speaker's informative intention (if he even has one) be made mutually manifest, not that the hearer actually recognize this intention. For S&W an assumption is manifest to someone merely because he is *capable* of representing it and of accepting its representation as true (p. 39); but then the thought of it might never occur to him and might never play a role in his thinking.

### Gricean theories

1. *Communicative intentions.* S&W agree with Griceans that communicative intentions are overt in a way that distinguishes them from other intentions, but they object that Gricean analy-

ses of such intentions are psychologically implausible. On such analyses these intentions either involve infinitely many sub-intentions or, if construed as reflexive, are, according to S&W, infinitely long. S&W neglect to mention an analysis which avoids this problem and which, in an article they cite, Grice (1969) himself suggests. On this account, what is required is not the presence of intentions at higher and higher levels but the absence of any “sneaky” intentions anywhere up the line. Moreover, S&W’s argument (pp. 256–57) that reflexive (or self-referential) intentions are infinitely long requires the premise that to refer to itself an intention (or any representation) must spell itself out, that is, include a mention of itself within itself (hence within itself within itself . . .). But consider the sentence T, “This sentence contains five words.” Its referent (itself) can be identified without even once substituting T for “This sentence.”

2. *Mutual knowledge/belief and the code model.* S&W assume that the only reason for “pragmatic” theories to impose a mutual knowledge/belief requirement on communication is adherence to the “code model.” In fact, pragmatic theories such as our own (Bach & Harnish 1979) are inferential, at least beyond the level of language comprehension (which is a matter of decoding). Indeed, we explicitly disavow the code model in our Introduction. Whereas for us the point of this requirement is to limit the speaker’s intention and the hearer’s inference to information that each can expect the other to rely on, S&W suppose that its point is to *guarantee* successful communication (p. 17). S&W rightly insist that there is no such guarantee, but we pragmatists don’t claim that there is. Our theory points out that attempted communication can and often does fail; indeed, successful communication does not entail mutual knowledge of, or even mutual belief in, its success (1979, p. 87).

3. *Ex post facto explanations.* S&W assert that Gricean accounts fail to exclude the possibility of several equally justified ways of taking a given utterance in a given context (p. 37). However, S&W do not show that Gricean accounts allow this possibility when communication *succeeds*. Of course, because communication can and often does fail, no theory should exclude this possibility altogether. After all, there are cases in which an audience reasonably takes an utterance the wrong way or doesn’t know which of several equally plausible ways to take it. That is why the very same model used for successful communication can be applied to unsuccessful communication. Our speech act schema (1979, chap. 4), which gives the general form of inference an audience is intended to make, provides a number of different places where things can go wrong. Among them are places for mutual contextual beliefs; for when the speaker and his audience are wrong about each other’s beliefs, communication is likely to fail.

4. *Linguistic communication.* S&W do not acknowledge any fundamental difference between linguistic communication and “ostensive” communication in general. It is our position (1979, p. 14) that in the linguistic case there prevails a *communicative presumption*, to the effect that when someone says something to somebody he does so with a certain identifiable communicative intention. The audience relies on this presumption and, as a result, need not ascertain *that* the speaker intends to be communicating. In nonlinguistic cases, however, one must generally infer that there is a communicative intention to be identified, as well as identify it.

5. *Speech acts and what is communicated.* S&W complain that, according to speech act theory, “the assignment of every utterance to a particular speech-act type is part of what is communicated and plays a necessary role in comprehension” (p. 244). Of course a theory’s taxonomy need not be used in practice by communicators, but, speaking for ourselves, we don’t claim that it does. Although we describe the hearer’s task as identifying the speech act being performed by the speaker, we also provide a taxonomy of speech-act types in terms of types of propositional attitudes being expressed. For example, to make a

request is to express a desire that the hearer perform a certain action; for the hearer to understand a request, accordingly, is for him to recognize the utterance as expressing such a desire. Curiously, this seems to be S&W’s view as well, for they remark, quite rightly, that a prediction can be understood without being explicitly classified by its hearers as a prediction, provided they take it as communicating an assumption (expressing a belief) “about a future event at least partly beyond [the speaker’s] control” (p. 245). Note that in our view a speaker can express an attitude without actually possessing it. This difference forms the basis for our distinction between communicative and other intentions. A communicative intention is a reflexive intention, in particular, one whose fulfillment consists in its recognition, and to recognize such an intention is to identify the attitude being expressed.

## Linguistic constraints on pragmatic interpretation: A reassessment of linguistic semantics

Diane Blakemore

Department of French (Linguistics), University of Southampton, Southampton SO9 5NH, England

The fact that some aspects of linguistic form do not contribute to the truth-conditional content of utterances is frequently acknowledged but very rarely explained. This is not, perhaps, surprising, given the range of expressions and constructions that convey nontruth-conditional meaning and the variety of effects to which they give rise. Moreover, until very recently one could camouflage the lack of progress by designating all such phenomena as “pragmatic,” the assumption being that someone would eventually provide a pragmatic theory.

And now, indeed, we do have a pragmatic theory that sheds light on the problem of nontruth-conditional meaning. This is not to say that the domain of this theory is the nontruth-conditional. For Sperber & Wilson (S&W) the fundamental distinction is between those aspects of meaning that are decoded according to linguistic rules and those worked out on the basis of the context and the assumption that the utterance is consistent with a general principle of communication. As they emphasize, the linguistically encoded meaning of an utterance underspecifies its truth-conditional content: It provides only a “blueprint” for a proposition, or, as S&W call it, a “logical form.” A complete proposition can be obtained only on the basis of contextual information and the assumption that the speaker has aimed at optimal relevance.<sup>1</sup>

Because linguistic meaning is not directly truth conditional, there is no reason to think that every aspect of linguistic meaning is definable in truth-conditional terms – and hence that the differences between the (a) and the (b) utterances in the following should not be specified in a theory of linguistic semantics:

- (1) (a) She’s well meaning but indiscriminate.
- (b) She’s well meaning and indiscriminate.
- (2) (a) He is an Englishman; he is, therefore, brave (Grice 1975a).
- (b) He is an Englishman; he is, after all, brave.
- (3) (a) It was Susan who read the article.
- (b) It was the article that Susan read.

This is not to say that the contribution made by *but*, *therefore*, *after all*, and the cleft construction can be explained in linguistic terms. Stalnaker (1974, p. 212) suggested that the function of certain linguistic constructions is to impose constraints on the contexts in which utterances containing them can occur. However, without an adequate account of the selection and role of the context in utterance interpretation, it is hard to see why

there should be such arbitrary links between linguistic form and pragmatic interpretation.

According to S&W, interpreting an utterance is not just a matter of identifying the proposition it expresses. Hearers are expected to integrate that proposition with their existing assumptions for the recovery of contextual effects. Indeed, the very act of communication gives hearers a guarantee that they will be able to derive adequate contextual effects. But accessing and using contextual assumptions involves a cost, and the hearer's aim of obtaining contextual effects is offset by the aim of minimizing processing costs. This means that it is in the interests of speakers who wish their communication to succeed that the required contextual assumptions be easily accessible to hearers. In some cases the speaker may have grounds for thinking that the necessary assumptions are already accessible to hearers, and hence that they will recover the optimally relevant interpretation of their own accord. In other cases, however, speakers may feel that they need to constrain the hearer's interpretation by making certain contextual assumptions accessible, thus ensuring correct context selection at minimal processing cost. That is, given the hearer's aim of optimizing relevance, the existence of structures and expressions that constrain the choice of context is to be expected.

To illustrate, consider the sequence in (4):

- (4) A: Susan's not coming today.  
 B: Tom's in town.

Although A's utterance provides the hearer with an immediately accessible context for the interpretation of B's, it is not clear where exactly the relevance of B's remark lies. It could be relevant as an explanation for the fact that Susan is not coming; it could be evidence for A's claim; it could be relevant as the specification of the implication of A's remark; it could be relevant as something that contrasts with the fact of Susan's not coming; or it could be an attempt to dismiss A's remark as irrelevant. In real conversation, however, the connection between the two remarks would not be left unspecified, and B would make his intentions clear, either by intonation or by the use of such "discourse connectives" as *you see*, *after all*, *so*, *however*, and *anyway*.

In S&W's framework computing the effect of a newly presented proposition crucially involves inference. The role of contextual assumptions is to act as premises which combine with the new information to enable hearers to add to or improve their evidence for existing assumptions. This means that an expression which imposes a constraint on the relevance of the proposition it introduces is a constraint on the inferential computations that proposition may enter into. For example, *therefore* (utterance 2a) specifies that the proposition it introduces must be interpreted as a conclusion, whereas *after all* (utterance 2b) indicates that the proposition it introduces is a premise. In both cases hearers are expected to supply the contextual assumptions that allow them to establish the prescribed connection. Words like *anyway* do not express inferential connections, but rather specify the property of the set of contextual assumptions which must be combined with the proposition expressed for the deduction of contextual effects.<sup>2</sup>

Notice that according to this account, the meanings of these expressions do not themselves contribute to the propositional representation whose relevance is being assessed. If this is right, then they cannot be part of the level of semantic representation that S&W call "logical form." This suggests a nonunitary theory of linguistic semantics. On the one hand, there is the study of logical form – that essentially conceptual theory which deals with the way in which elements of linguistic structure map onto constituents of propositional representations. On the other hand, there is the study of linguistic constraints on relevance – that essentially procedural theory which deals with the way in which elements of linguistic structure affect pragmatic computa-

tions. Given the foundation provided by relevance theory, both promise to be fruitful areas for future research.

NOTES

1. For further discussion see the accompanying commentary by Carston.
2. For a more detailed analysis of these expressions see Blakemore (1987).

Being explicit

Robyn Carston

Department of Linguistics, University College London, London WC1E 6BT, England

It is a truism that verbal communication is one of the primary bases for human beings' understanding of each other and for their establishing relationships with each other. As participants in this practice we accumulate a range of common experiences. For example, most of us find exchange with some people dreadfully tedious, whereas with others we feel stimulated, challenged, or drawn closer. We find some conversationalists witty or amusing, others long-winded or patronising. Sometimes the act of uttering is much more important than the actual content of the utterance, just as in other contexts the fact of silence is highly significant. Speakers may betray assumptions and attitudes they haven't intended to convey; hearers may experience a nagging uncertainty about whether a speaker is making some insinuation or whether they are themselves responsible for projecting it into the interpretation. One could multiply this list of familiar observations. With the advent of relevance theory, however, we need not merely note these facts and vaguely wonder what they indicate about human psychology. Sperber & Wilson's (S&W's) *Relevance* presents a theory that enables a systematic and explanatory account of such experiences and sets up a programme for sustained research into these and many other interpretative issues. It is important to keep in mind that what underlies the relevance-theoretic treatment of verbal communication is a theory of communication in general that is based in turn on a theory of human cognition. As such, relevance theory is the first account of pragmatics which is grounded in psychology.

In discussing the meaning of utterances Grice (1975a; 1978) distinguished what is said from what is implicated. He took what is said (that is, the explicit content of the utterance) to be the proposition determined by three factors: the conventional meaning of the linguistic form of the utterance, the assignment of referents to referring expressions, and the resolution of any ambiguities of sense. However, in demonstrating his various pragmatic maxims at work, he ignored explicit content and concentrated entirely on implicitly conveyed propositions, implicatures. These do show very clearly how the utterance of a certain linguistic form in a particular context can convey much more than that form itself literally does. For example:

- (1) A: Will Mary come to the movie with us?  
 B: She has an exam in the morning.

Assuming that some maxim of rational communication is in operation, in the context of A's question B's utterance implies that Mary will not or should not go to the movie. But of course the speaker hasn't actually said this: The meaning of the linguistic form used does not include anything about movie-going, and the truth or falsity of B's utterance is independent of whether or not Mary goes to the movie.

Now linguists such as Gazdar (1979), Horn (1984), and Levinson (1983), who took up Grice's principles, have followed his example and been almost entirely concerned to use Gricean principles to account for the generation of these sorts of implica-



tions. They have written as if any meaning not derived by linguistic decoding must be implicated. However, this can not be right, since it would entail that the results of reference assignment and disambiguation are also implicatures. In fact, as soon as one considers actual utterances it is obvious that pragmatic principles must play a large role in determining explicit content because utterances are frequently fragmentary or vague and may be ungrammatical or contain slips of the tongue. Interpretation involves deriving the framework or schema (a logical form) provided by linguistic content but it is seldom, if ever, exhausted by it.

S&W define the explicit content of an utterance: "An assumption communicated by an utterance *U* is *explicit* if and only if it is a development of a logical form encoded by *U*" (p. 182). Any assumptions explicitly communicated are a combination of linguistically encoded and contextually inferred features. There is always a linguistic contribution, but its role varies along a continuum from near total determination of the explicit content to the merest indication, as shown by the following examples of possible utterances:

2. The sun will rise at 5:25 am on May 15, 1989.
3. Susan's playing is good enough.
4. He admitted it.
5. At home.
6. Soon.

Understanding these requires various amounts of disambiguation, reference assignment, enrichment (e.g., in (3) "good enough" for what?), and completion. These processes involve an interaction of the linguistically derived meaning with contextual assumptions and are driven by the bid for maximal relevance and constrained by the principle of relevance.

Philosophers of language have a longstanding interest in the truth-conditional content of sentences and utterances. So they have not confined their interest to linguistic content alone but, like Grice, have assumed that determining explicit content involves the fixing of certain indexical expressions by the context so as to arrive at a propositional form, that is, a form capable of being true or false. However, this does not conform to the reality of psychological processes either, as S&W show in the following example:

7. Repairing this watch is going to take some time (p. 189).

The interpretation recoverable from this utterance by decoding the linguistic content and assigning reference to "this watch" will give a complete proposition expressing the simple truism that the repairing of the watch will be an activity taking place over time. This is truth-evaluable; it will be true in almost every situation one can conceive of. However, recall that according to the principle of relevance the interpretation chosen is the most accessible interpretation (the one requiring least effort) which gives sufficient cognitive or contextual effects. On this basis, the truism is rejected as the explicit propositional content the speaker wished to convey because it is not sufficiently relevant: It does not interact with contextual assumptions to yield an adequate range of contextual effects. It must be enriched in some way in order to do so; the most probable enrichment, in the sort of context likely to prevail for the customer wanting her watch repaired, is that the time period involved will be longer than she might have expected, an assumption bound to have a range of implications for her. It seems that the formal philosophical requirement of minimal truth evaluability is an arbitrary principle with no force in a psychologically adequate account of communication.

The relevance-theoretic account of explicit content opens the way for a solution to a number of problem cases that motivated Grice's original postulation of pragmatic maxims, including the apparent ambiguity of logical connectives such as *and* (truth-functional conjunction with possible temporal and/or causal connotations) and *or* (inclusive and exclusive), and of the num-

bers (an "at least" sense and an "exactly" sense). I argue (Carston, forthcoming) that although none of these involves linguistic ambiguity of the word itself, the correct level for the incorporation of these pragmatically derived aspects of meaning is the level of explicit content rather than that of implicature. So, for example, the causal connection between the conjuncts of (8) is part of the assumption explicitly communicated, even though it is neither given by linguistic content nor required for the derivation of a truth-evaluable proposition.

8. He was hit by a bus and died instantly.

This analysis follows directly from the principle of relevance which directs the recovery of assumptions, whether explicit or implicit, that have a satisfactory range of effects for the least effort in deriving the effects.

Pragmatics has been taken up by linguists and philosophers and generally suffers from the preoccupations of their disciplines. Relevance theory, on the other hand, gives a cognitive psychological basis to what surely is a psychological phenomenon: human beings' interpretation of each other's utterances.

## Relevance to what?

Herbert H. Clark

*Department of Psychology, Stanford University, Stanford, Calif. 94305*

The achievement of relevance by people in conversation is an extraordinary feat and well worth studying. So is the achievement of *Relevance* by traditional adversaries from opposite shores of the English Channel. Sperber & Wilson (S&W), burying their Gallic and Celtic hatchets, have offered us a provocative account of communication and cognition. They are to be congratulated on it. Not that the account isn't flawed, but what account isn't? In two earlier papers (Clark 1982; Clark & Gerrig 1984), Gerrig and I have argued against S&W's earlier views on mutual belief and irony and, despite their protestations, our arguments still stand. Here I will look at two more general features of their enterprise.

**Collective Action.** The centerpiece of *Relevance* is, of course, relevance. According to the OED, *relevant* means "bearing upon, connected with, pertinent to, the matter in hand." But for S&W, what is the matter in hand? To what purpose is a communicative act relevant? For all their talk about relevance, they never really say. That, I suggest, jeopardizes their entire enterprise.

It was Grice's (1975a; 1978) and Austin's (1962) insight that communication must be viewed as part of a theory of action – of what people do. At first, S&W appear to follow suit as they discuss informative and communicative intentions and ostensive-inferential communication. But they stop short of embracing Grice's and Austin's full insight. They are satisfied in claiming that "The aim of communication in general . . . is to increase the mutuality of cognitive environments and thereby the similarity of thoughts" (Précis, sect. 4.2, para. 5). Grice and Austin might accept this as the *means* of communication, but surely not its *aim*.

People talk, Grice (1975a) and Austin (1962) tacitly assumed, as a means of doing things they can only do collectively – arguing, instructing, negotiating business, performing rituals, telling stories, gossiping, planning. This is really the point of Grice's observation that talk exchanges "are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction" (Grice 1975a, p. 45). For Grice, this evolving common purpose is the basis for all conversational implicatures. In hindsight, we can see that it was a mistake for him not to develop the notion any

further. What he offered instead were rules of thumb, or maxims, by which one could determine the effects of common purpose. But these rules of thumb are no more than that. They are proxies, promissory notes, for a theory of collective action that is yet to be developed.

S&W, however, spurn Grice's observation and decline to develop any notion of "matter in hand" or "common purpose." When they characterize *relevance*, it is always divorced from what the participants in a discourse are really doing. "For us," say S&W (p. 161), "the only purpose that a genuine communicator and a willing audience necessarily have in common is . . . to have the communicator's informative intention recognised by the audience." Perhaps. But even that intention cannot be recognized without seeing the potential common purposes to which the communicator's action is to contribute. S&W's position is like claiming that the only purpose I need in stepping on the car's accelerator is to put more gasoline into the carburetor. I do feed the carburetor, but that hardly accounts for why I usually take that action. I do it to speed up the engine, to turn the wheels faster, to speed up the car, to get me to my destination quicker, and so on. The "matter in hand," the main purpose for an action, is often quite remote from its immediate effects. Likewise in communication. That was Grice's and Austin's shared insight. So what S&W leave us with is a peculiarly empty notion of relevance. It almost belies the title of the book.

As a result, S&W pass off onto "cognitive psychology," without further explication, what for many scholars are the central issues of pragmatics. Consider their idea that addressees take the decoded content of an utterance along with what is "mutually manifest," weigh its "contextual effects" against its "processing effort," and select the interpretation that is "optimally relevant – that is, "the first accessible interpretation consistent with the principle [of relevance]" (Précis, sect. 3.3, para. 6). But what exactly are "accessibility" and "processing effort"? Do we really *select* interpretations, as they presuppose? How are contextual effects weighed against processing effort? And so on. Paradoxically, to answer these questions, we would need just the notions of collective action and evolving purpose that S&W are unwilling to provide.

Notions like these are already under investigation in the literature on conversation and other types of discourse, but S&W pass them over. In the end, however, they will not be able to duck the question "relevant to what?"

**Layers of communication.** Many types of discourse have more than one distinct layer of action or communication (Bruce 1981; Clark 1987; Goffman 1974). S&W, however, presuppose that all communication is flat – that it has only one layer, one type of relevance. That, I suggest, leads to a misrepresentation of many important phenomena in communication.

Layering is easiest to recognize in fiction – novels, plays, short stories, jokes, films – though it is common elsewhere too. In Herman Melville's novel *Moby Dick*, the narrator Ishmael is telling some landmen, perhaps in a Nantucket tavern, about his whaling adventures. Call this domain 1. Yet everything in domain 1 – Ishmael, his audience, his adventures, *and* his narrative – are Melville's inventions, and in writing the novel, he "communicated" that narrative to us readers. Call this domain 2. (One can argue for another domain between 1 and 2, but it is not needed to make the point.) Likewise, in *Hamlet*, Hamlet talks with Ophelia in a fictional domain 1, but Shakespeare "communicates" that conversation to us playgoers in domain 2 (via actors in yet another domain pretending to be Hamlet and Ophelia). All fiction has at least two domains, two layers.

How does relevance theory apply? In domain 1, the theory might claim that, fictionally, when Ishmael says, "Call me Ishmael," he has "informative" and "communicative intentions" toward his Nantucket audience, and they in turn presume that his utterance is relevant to them. The theory might also apply when Hamlet tells Ophelia, "Get thee to a nunnery."

What about domain 2? When Melville writes for us, "Call me Ishmael," *we* aren't to call *him* Ishmael. Nor is Shakespeare asking us (or Ophelia or the actress playing Ophelia) to hie off to some "nunnery" – some brothel. Domain 2 is somehow very different from domain 1.

Relevance theory doesn't go beyond the surface layer. It has nothing to say about domain 2, even if we avail ourselves of S&W's notion of "interpretation" or "interpretive representation." Here is why. Melville and Shakespeare have intentions toward us, but these are not "informative" or "communicative intentions" – they do not constitute the Gricean speaker's meaning (Clark 1987) – and the principle of relevance does not apply. So when Melville and Shakespeare "communicate" with us, it is communication of a fundamentally different type. This has a surprising but demonstrable consequence. Even if relevance theory could explain how the Nantucket landmen understood Ishmael, and how Ophelia understood Hamlet, it would not explain how *we* do. Relevance theory simply does not apply to a great deal of our most cherished communication.

Relevance theory has a long way to go to become a full theory of communication and cognition. It cannot work, I suggest, without well-developed notions of collective action and layering. But can it accommodate these without being stretched beyond the breaking point? That may be the next test of the trans-Channel alliance.

#### ACKNOWLEDGMENTS

This work is supported in part by NSF Grant BNS 83-20284. I thank Eve V. Clark, Florence Edwards, and Ellen A. Isaacs for comments.

### The task of the speaker and the task of the hearer

Anne Cutler

MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EK, England

These brief remarks will be addressed to Sperber & Wilson's (S&W's) view of verbal communication. First, S&W draw a distinction between two separate processes of comprehension: a decoding process and an inferential process. They are principally concerned with the operation of the latter; the former they dismiss as automatic and therefore "not so much a part of the comprehension process as something that precedes the real work of understanding" (p. 177). Second, they imply that the work of understanding (though "real") is less than the work of speaking; the brunt of the work in communication is borne by the speaker. "It is left to the communicator to make correct assumptions about the codes and contextual information that the audience will have. . . . The responsibility for avoiding misunderstandings also lies with the speaker" (p. 43). "If the speaker has done her job properly, the end of the utterance should confirm all the provisional choices . . . that have been made along the route" (p. 208).

This picture is distinctly unfair to the hard-pressed hearer. Hearers are presented with signals which are for the most part semantically, syntactically, lexically, and phonologically unpredictable; moreover, the signals arrive in a noisy channel and are frequently subject to considerable distortion and attenuation. Speakers, on the other hand, have in principle a free hand in what they choose to say and how they choose to say it. S&W's principle of relevance is based on the observation that speakers do not take advantage of this freedom; in contrast, they constrain their utterances quite severely in order to make life easier for hearers.

In fact, S&W have here revealed only the tip of an iceberg. Speakers construct their speech output so as to cater to listeners' needs in a far more detailed fashion than is captured by the

guarantee of relevance or by Grice's injunctions to speakers not to bore, puzzle, offend, or deceive audiences. In particular, there is abundant evidence that speakers adjust their output to assist the listener at those levels which S&W claim are the subject of "automatic" processing – even at the level of segment production, as the following examples will show.

On the one hand, consider the inhibition of certain phonological rules of elision and assimilation. The application of such rules can result in a distortion, in casual speech, of phonetic segments which would be clearly articulated in more formal speech (Cooper & Paccia-Cooper 1980; Kaisse 1985); for example, the sequence [tj] can become the affricated segment [tʃ]. This palatalisation rule can apply across word boundaries, as in "Meetcha after work?" Cooper and Paccia-Cooper investigated the applicability of such palatalisation as a function of the informativeness of words preceding and following the boundary. For example, they varied word frequency of occurrence, comparing relatively common words ("rode your horse"; "had utensils") with much less frequent ones ("goad your horse"; "had euglena"). Varying the frequency of the word preceding the boundary had no effect on the frequency of palatalisation across the boundary; but varying the frequency of the word after the boundary had a strong effect – palatalisation was used significantly less often before rare words. Cooper and Paccia-Cooper also looked at the effect of contrastively stressing each word; again, stressing preceding words had no significant effect on the applicability of palatalisation, but stressing following words almost completely inhibited it.

In other words, distorting the ends of words is something speakers are fairly happy to do; but they are reluctant to distort word beginnings if the words are either rare or contrastively stressed, that is, if their information value is high. The beginning is the most important part of a word for the listener – distortion of word onsets disrupts word recognition far more than distortion of later segments (Bagley 1900; Cole 1973; Marslen-Wilson & Welsh 1978). So the speakers in Cooper and Paccia-Cooper's studies were clearly making phonological choices in such a way as to minimise disruption to the listener.

The same kind of motivation can be discerned in a pattern observed by Cutler (1983) in the correction of slips of the tongue. Errors of lexical stress occur quite frequently – *synTAX* for *SYNTAX*, *orIgin* for *ORIGIN*. Mostly such errors remain uncorrected by the speaker. This should cause the hearer little problem, since prosodic stress plays no role in word recognition (Cutler 1986); the hearer will probably notice a mismatch between spoken form and canonical lexical form, but will be readily able to discount it (cf. p. 23). What does disrupt word recognition, though, is getting vowel quality wrong – substituting a full for a reduced vowel or vice versa (Bond & Small 1983). So it is not surprising to find that precisely those stress errors which result in a change of vowel quality are the stress errors most likely to be corrected. Thus *orIgin*, in which a full vowel in the initial syllable has been replaced by a reduced vowel, and a reduced vowel in the second syllable has been replaced by a full vowel, is much more likely to be corrected by the speaker than *synTAX*, in which both vowels are full in both target and error.

These segment-level instances of perceptually driven speaking are striking; but one could easily add many instances at the lexical level (such as the tendency for nonce formations not to distort the real words on which they are based; Cutler 1980) or at the prosodic level (such as the fact that the greater the semantic contrast between a lexical slip of the tongue and the intended word, the more likely it is that the speaker will draw hearers' attention to a correction of that slip by stressing it; Levelt & Cutler 1983). Seen in this light, speakers' attention to ensuring relevance is merely one end of a continuum of hearer-coddling; there is certainly nothing special about it, and nothing that makes attention to hearers' inference processes qualitatively different from attention to hearers' decoding processes.

Thus, there is a sense in which the task of the speaker extends beyond the translation of a message into a spoken output;

speakers take upon themselves some responsibility for ensuring that hearers successfully accomplish understanding. But they do this purely out of self-interest, to ensure that their message gets across, and they do it precisely because the task of the listener is intrinsically so much harder than the task of the speaker. Particularly, it is harder at exactly those levels which Sperber & Wilson dismiss as the province of automatic and reflex processing. At these levels speakers strive to ensure reception of their message by hearers. Decoding is part of the work of understanding too.

## Relevance and mutual knowledge

Martin Davies

*Philosophy Department, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England*

It is common for philosophers of language to abstract from considerations of context dependence. The resulting picture of language is an idealised one; but the assumption behind much work over the last twenty years is that the idealisation does no serious theoretical damage. This assumption is, in turn, nourished by the idea that context dependence in general can be treated on the model of the indexicality exhibited by the word "I". The meaning of "I" specifies a very simple rule for assigning a reference relative to a context. Give or take "I," "here," "now," and a few other expressions to be treated on the same lines, the linguistic meaning of a sentence will determine a proposition expressed and, in particular, will determine truth conditions.

One of the major virtues of Sperber & Wilson's (S&W's) *Relevance* is its stress upon the fact that the common picture is not just idealised but mythological; truth conditions are radically underdetermined by linguistic meaning.

Whether or not the bold claims for relevance theory can be sustained at every point, the treatment of metaphor in *Relevance* and the subsequent paper (Sperber & Wilson 1986b) is a real advance. It shows that the apparent dichotomy between speaker-meaning accounts of metaphor following Searle (1979) and "seeing as" accounts following Davidson (1978) is spurious, thus correcting an impression given, for example, by Davies (1983). And it gives some determinate theoretical substance to the suggestion of Blackburn (1984, pp. 171–79) that a metaphor is an "invitation to explore" a comparison or image.

The radical underdetermination of truth conditions by linguistic meaning is enough to show that something is seriously wrong with the code model of communication: Even in the case in which there is only a single determinate thought to be communicated, the content of the thought is not fully encoded in the sentence uttered. But there is something else wrong with the code model as a model of human communication – that is, communication amongst creatures for whom there is a difference between entertaining a proposition, or having a proposition presented as a candidate for belief, on the one hand, and actually going forward in judgement and believing the proposition, on the other.

If propositions really were encoded in sentences, and I were equipped with a mechanism for decoding, then upon hearing an utterance of a sentence I would be presented with a proposition as a candidate for belief. This would be similar to the way in which, if I have a perceptual experience of the world as being a certain way, then the proposition that the world is that way is presented as a candidate for belief. In the case of perception, if I take my experience at face value, then I believe that the world is indeed that way. What is more, I usually do take my experience at face value. I do not require a justification for doing that; rather, I should need a reason not to take it so (cf. pp. 257–58, n. 28). But still, the difference between perception and belief remains. In the case of communication on the code model, it



may also be that I usually take what I hear at face value and would require a reason not to do so. Yet the difference between decoding and belief remains.

As a theory of human communication, the code model cannot be adequate by itself. In order even to be a candidate, it would need to be augmented by a component reflecting the sensitivity of beliefs to reasons. One kind of augmented code model would make use of the notion of mutual knowledge. Communication would be characterised as depending upon a shared code plus mutual knowledge between speaker and hearer that the code is shared – along with assumptions on the hearer's part about the sincerity and reliability of the speaker.

Many philosophers object to this kind of augmented code model on the grounds that it represents the ordinary communicative use of language as far more reflective, calculating, and psychologically convoluted than it actually is. Sometimes (McDowell 1980) this line of objection manifests itself in the suggestion that human communication is a mode of instilling information importantly similar to animal communication. However – as Strawson (1980) in effect points out – to assimilate human communication to animal communication would be to ignore the crucial sensitivity of beliefs to reasons. If the hearer doubts that the speaker is sincere and reliable, or – more important here – doubts that the speaker and the hearer share a common code, then the rational hearer will refrain from taking what he hears at face value, just as the rational perceiver may sometimes refrain from taking his experience at face value.

What is needed in an augmented code model is, apparently, a condition about the absence of certain undermining doubts rather than about the presence of certain justifying beliefs. This prompts the idea that the philosophical work which was to be done by the notion of mutual knowledge should instead be assigned to a negatively characterised notion: mutual absence of doubt. For example, neither speaker nor hearer doubts that they share a code, neither party thinks that either party doubts that they share a code, neither party thinks that either party thinks that either party doubts that they share a code, and so on. Rather than the psychologically implausible claim that in communication indefinitely many beliefs must be present, we have here the psychologically harmless claim that indefinitely many doubts must be absent.

S&W have their own replacement for mutual knowledge in the notion of mutual manifestness (p. 42), though they do not, of course, use that notion to rehabilitate the code model. A claim of mutual manifestness unfolds into an infinite series of manifestness claims, each of which is a claim that an individual is capable of mentally representing a state of affairs. The truth of an infinite list of manifestness claims does not, of course, require that anyone should be capable of simultaneously mentally representing infinitely many states of affairs. So mutual manifestness escapes a claim of psychological implausibility analogous to that leveled against mutual knowledge.

One might wonder whether it is really true that subjects are capable of each of the very complex pieces of mental representation assumed by the theory. But that is a minor quibble. A more interesting question is this: To what extent could mutual absence of doubt do the various pieces of work for which mutual manifestness is cast in the theory of relevance?

It may appear bizarre that I have introduced this question via augmentation of the code model, given that the code model is clearly wrong on other grounds. But it is certainly a part of S&W's view that human communication often exploits the code of language. Now if a speaker exploits a code, then his intentions towards his hearer are rationally sensitive to the absence of certain doubts. The speaker does not think that he fails to share a code with the hearer. The speaker does not think that the hearer thinks that they fail to share a code. And so on. Consequently, my second question is this: Can we give an account of what it is for a speaker to exploit a code without invoking some notion like mutual absence of doubt?

Sperber & Wilson say this: In discussing relevance and

ostension (p. 50), "Ostensive behaviour provides evidence of one's thoughts. It succeeds in doing so because it implies a guarantee of relevance. It implies such a guarantee because humans automatically turn their attention to what seems most relevant to them." The argument is very compressed. It begins from a general claim about humans as information-processing systems: namely, that the system is geared for efficiency. In particular, the natural mechanisms for directing attention generally subserve the efficient gathering of information. It is quite consistent with this general claim that, on specific occasions, attention may be attracted to what is in fact not a source of useful information.

From the general claim about the natural order it is said to follow that drawing someone's attention to a phenomenon "implies a guarantee of relevance." How exactly does this argument go? Suppose that the natural function of attention is to provide useful information, and that I intentionally seize or direct someone's attention. Then, we might say, the natural normative ideal would be that the person will gain some useful information. There is, in that sense, a natural guarantee of relevance.

But this natural guarantee of relevance is not clearly the right sort of thing to support the conclusion of the argument – namely, that my intentional direction of a person's attention provides that person with evidence about my intentions. One thing that would mediate the provision of such evidence would be a standing presumption that the social order measures up to the natural order in respect of the connection between attention and information.

The principle of relevance says that each act of ostensive communication communicates the presumption of its own optimal relevance. I am not questioning that principle. What I am asking, as my third question, is this: Can the principle of relevance be explained in terms of the claim that human beings are naturally efficient processors of information?

## Relevance theory, mutual knowledge, and accidental irrelevance

Richard J. Gerrig

Department of Psychology, Yale University, New Haven, Conn. 06520

If Sperber & Wilson (S&W) wish to banish mutual knowledge from accounts of language processing, they must demonstrate that relevance theory can accommodate the range of phenomena for which mutual knowledge has traditionally been invoked. Consider, for example, their brief treatment of the ambiguous referring word "Jennifer" in the utterance, "Jennifer admitted stealing" (*Relevance*, p. 206). They create a situation in which the (male) addressee "knows a Jennifer Smith and a Jennifer O'Hara," and they hypothesize that the addressee "proceeds to make a tentative assignment of reference to the expression 'Jennifer' by considering whether the information that Jennifer Smith did something or the information that Jennifer O'Hara did something might be relevant to him in some context he currently has accessible." Absent in this analysis is any consideration of the importance of the speaker's identity in disambiguating this type of referring phrase. For example, if the sentence "Jennifer admitted stealing" were uttered by Tim O'Hara, whom the addressee knew to be Jennifer O'Hara's husband, we would expect that in most contexts *relevance* would have little effect on the interpretation of "Jennifer." This expectation of relevance's limited role – which extends to all situations in which some uniquely salient "Jennifer" is mutually known to a speaker and addressee – is bolstered by cases of *accidental irrelevance*.

To achieve accidental irrelevance, we must manipulate the knowledge states of the speaker and addressee. Suppose, for example, that Tim O'Hara had uttered "Jennifer admitted stealing" to his uncle, Martin. Suppose, further, that Martin

knew with great certainty that Jennifer had admitted to stealing even before Tim's utterance, a fact of which Tim was unaware. According to S&W, an utterance that replicates information already known with certainty by an individual has no contextual effects and, consequently, no relevance. Because the interpretation of Jennifer as "Jennifer O'Hara" would thus make the utterance irrelevant, Martin would have to seek another interpretation: If S&W are right, the search for relevance would inevitably cause Martin to attribute the admission of stealing to some inappropriate Jennifer. Uncle Martin might, for example, infer that Jennifer Smith was the confessed criminal, then proceed to deduce whatever contextual implications would follow from that assumption. This is not an isolated example. Relevance theory appears to predict that addressees would be led astray on any occasion in which a speaker (accidentally) uttered information of which the addressee was previously certain.

If S&W will allow mutual knowledge back into their formulation, this difficulty can be eliminated. Note again that in my scenario Tim spoke with no intention of violating Martin's presumption of optimal relevance: He simply did not know that he and Martin already shared the information about Jennifer. Thus, although Tim's utterance is irrelevant relative to the entire universe of contexts Martin might have accessible, it is fully relevant within the context defined by the information that is mutually known. My suggestion is that S&W must precede their statement of the *principle of relevance* with something like, "Based on the communicator's assessment of mutual knowledge." What this addition acknowledges is that we cannot assure relevance in any absolute sense – because our addressees will have acquired a great deal of information outside our purview. We can only assure relevance relative to our mutual knowledge base.

S&W are right to argue that the requirements for mutual knowledge as set forth in philosophy are psychologically unrealizable: As finite devices we are incapable of assuring ourselves of the truth of an infinite series of assumptions. Fortunately, Clark and Marshall (1981; see also Gibbs 1987) have suggested a heuristic approach to the assessment of mutual knowledge that is psychologically plausible. From Lewis (1969), they adapt the *Mutual knowledge induction schema*:

- A and B mutually know that *p* if and only if some state of affairs *G* holds such that:
1. A and B have reason to believe that *G* holds.
  2. *G* indicates to A and B that each has reason to believe that *G* holds.
  3. *G* indicates to A and B that *p*. (Clark & Marshall 1981, p. 33).

As Clark & Marshall put it, "the point of this schema is that [A] and [B] don't have to confirm any of the infinity of conditions in mutual knowledge at all. They need only be confident that they have a proper basis *G*, grounds that satisfy all three requirements of the induction schema" (pp. 33–34). Clark and Marshall go on to describe a series of heuristics of *copresence* that enables speakers and listeners to be certain that such grounds exist.

Without an acknowledgement that speakers can be only as relevant as their mutual knowledge allows, relevance theory makes the counterintuitive prediction that addressees will make highly inappropriate inferences in (at least) the situations in which speakers have accidentally reported old information. Clark and Marshall's psychologically plausible account of mutual knowledge should encourage Sperber & Wilson to remedy this aspect of their theory.

## The relevance of Relevance for psychological theory

Raymond W. Gibbs, Jr.

Program in Experimental Psychology, University of California, Santa Cruz, Calif. 95064

Sperber & Wilson (S&W) have provided a rich set of hypotheses for a pragmatic theory of utterance interpretation. Their main thesis – that utterance interpretation can work without an appeal to some notion of mutual knowledge – is an interesting departure from most contemporary pragmatic theories which, if true, would impose significant constraints on psychological theories of language processing. My contention is that there are serious difficulties with *any* psychological theory of utterance interpretation which does not include an evaluation of speakers' and listeners' mutual beliefs as part of the process of recognizing what speakers mean. I will restrict my comments here to three points which, I believe, are problematic for S&W's account of utterance interpretation (see Gibbs, 1987, for an elaboration of these remarks).

**Mutually known versus mutually manifest.** The suggestion that utterance interpretation can be explained without mutual knowledge, but via some appeal to the idea of mutual cognitive environments along with the *principle of relevance* might, at first glance, seem to be a reasonable one. S&W argue that problems inherent in establishing mutual knowledge between speaker and listener may be so great as to be practically insoluble. However, their approach suffers from some of the same difficulties they wish to remedy. After all, if there is a problem establishing some knowledge or beliefs as being mutually *known*, then there are likely to be similar problems in recognizing that some cognitive environments are mutually *manifest*. S&W argue, nonetheless, that the notion of what is manifest to an individual is weaker than the idea of something that is actually known or assumed and that therefore, something can be manifest without really being known.

The issue of what kinds of experience constitute "knowing" or "assuming" as opposed to merely "being manifest" is a delicate one. Yet making a distinction between something being "known" and other things being "manifest" may artificially create a difference where none exists. Speakers and listeners possess *tacit* knowledge which is difficult to access consciously, as in our knowledge of grammar. But we still think of our knowledge of grammar as "knowledge," no matter how tacit it may seem. Even though this knowledge is difficult to specify as a set of mutually held propositions, this does not imply that speakers and listeners are unable to mutually recognize the existence of the shared knowledge. Listeners do not have to access a relevant assumption *consciously* before a speaker can say something that makes it possible to comprehend an utterance, as long as it is mutually known to both parties that the listener is capable of using this tacitly shared information at the right moment. It appears that S&W are "sneaking" mutual knowledge in the back-door of their theory and have adopted a framework for describing verbal communication which depends crucially on the very concept they wish to abandon.

**The necessity of mutual knowledge.** There are many instances of language in which the speaker and listener must share certain knowledge and beliefs if an utterance is to be correctly understood. Consider the following exchange:

Rick: Are you going to the party tonight?

Alice: I hear Jack's coming.

Successful interpretation of Alice's response demands that the listener make an inference about what she means. But recognition of the speaker's intention is of a special kind, what Grice (1968) called an *m-intention*. An *m-intention* is a speaker's intention to produce an effect in the listener by means of the hearer's recognition of that intention. Thus, Alice wants Rick to recognize her intention in part by means of Rick's recognition of

it. This requires that Rick have not only some knowledge about Alice's attitudes toward Jack, but also some idea that Alice believes that Rick knows this about her. The reason Alice says what she does, and does not make a more direct response to the question, is that she intends Rick to base his inference not just on *any* knowledge of beliefs he has, but on their *mutual* knowledge and beliefs. If Rick successfully recognizes this intention, he will draw the "authorized" inference (Clark 1977a). S&W's proposal that understanding Alice's response requires the listener, Rick, to find a context which makes Alice's assertion most relevant does not sufficiently constrain the interpretation process and would lead to the generation of both "authorized" and "unauthorized" inferences.

A similar difficulty lies in the interpretation of irony. Consider the following conversation.

Bob: It's a lovely day for a picnic.

(They go for a picnic and it rains.)

Mary: It's a lovely day for a picnic, indeed.

S&W propose that Mary's utterance will be seen as ironic because it echoes Bob's previous statement, a fact that Bob must recognize if he is to understand what Mary means. To do this, however, listeners such as Bob must share enough information to be able to recover the source of what is echoed. A speaker such as Mary would not produce an echo unless she assumed that Bob shared the right knowledge to locate the source of the echo and thus to see her utterance as ironic. In fact, it is the speakers' knowledge of who does and who does *not* share certain information that makes it possible to have "victims" of sarcasm or irony. A speaker at a social gathering might overtly compliment an addressee by saying "Your dress is just perfect," yet intend overhearers to recognize the remark as sarcastic because they share the particular belief with the speaker that the addressee's attire is inappropriate for the occasion, unbeknown to the "victim." It is precisely *because* speakers and listeners mutually know certain beliefs that sarcasm has its special bite. What speakers and listeners presuppose about each other's knowledge and beliefs permits the successful use of sarcasm in conversation.

**Relevance theory as process model.** S&W suggest that a trade-off exists during utterance interpretation between maximizing contextual implications and minimizing processing effort. Presumably listeners derive the literal, context-free proposition of an utterance before they choose a context in which that proposition is viewed as most relevant. The listener relies on background knowledge, but there need be no assumption that this knowledge is mutually *known* beforehand. This kind of process model would predict that understanding utterances whose literal meanings differ from their intended interpretations requires additional processing effort. The psycholinguistic evidence strongly indicates, however, that this is not the case. Much research has shown that listeners need *not* analyze the literal meanings of many types of figurative utterances, such as indirect speech acts, metaphor, sarcasm, and idioms (see Gibbs 1984; 1986 for reviews), before determining their conveyed interpretations. If people do not automatically analyze the literal meanings of utterances, then S&W are wrong in assuming that conversational inferences (conveyed meanings) can be determined by finding a contextual assumption which makes some proposition (or "literal meaning") most relevant.

These brief comments highlight some of the difficulties I see for S&W's theory. They have written an interesting, courageous book, but it would be premature to accept their theory as a psychological model of linguistic processing.

## Relevance theory and the scope of the grammar

Liliane Haegeman

Department of English, University of Geneva, 1205 Geneva, Switzerland

Apart from the obvious significance of Sperber & Wilson's (S&W's) book for domains traditionally called "pragmatics," "discourse analysis," "text grammar," and so forth, research in syntax too will be profoundly affected by the implications of relevance theory. My commentary focuses solely on this aspect of their work.

Anyone familiar with work in theoretical syntax must at times be disappointed by the cavalier way in which "pragmatics" is treated by syntacticians who otherwise insist on and apply the most rigorous principles in their research.

Until recently, syntacticians might have answered this criticism by arguing that it is nowhere near obvious that there exists a coherent theory of utterance interpretation anyway. S&W's book puts an end to this argument. Failure to admit to the existence of a theory such as S&W's may entail that the syntactician include in his syntactic analysis phenomena which would far better be treated in terms of utterance processing. I shall illustrate this point with an example from the literature.

Recently some authors have discussed in more or less detail the phenomenon of indefinite object drop in English as exemplified in (1):

- (1) John is eating.

In her discussion of such examples Zubizarreta (1985, p. 250) says:

It is possible to say both *The baby is eating cereal* and *The baby is eating a marble* but the intransitive sentence *The baby is eating* necessarily implies that the baby is eating some sort of food, not a marble. The intransitive verb *eat* therefore takes as an "inherent object" the constant semantic argument "food."

Similarly, Rizzi (1986, n. 5) argues that the missing object in (1) is somehow interpreted as canonical or prototypical. He suggests that this is a lexical property of the nontransitive *eat*.

Both authors cited assume that the obvious interpretation of (1) – that is, John is eating food or a meal – ought to be made part of the semantic, grammar-based representation of (1). However, assuming a theory of utterance interpretation such as that proposed by S&W, this is by no means the only possible analysis. I shall sketch an alternative way of handling the data here.

In the unmarked case (1) will be taken to mean that John is eating a meal; but note that in a context in which John has been unable to eat for a certain time, (1) may merely refer to John's taking any food at all.

In more exotic contexts the interpretation of the implicit object may even be extended. For example, a researcher in a medical context who is trying to study the result of eating marbles might well refer to a subject of his experiments who is finally complying with the imposed test with (1). It is the specific context outlined here that will trigger the rather less obvious interpretation. This observation makes it plausible that the "normal interpretation" of (1) is likewise a matter of contextualization rather than of grammar-internal or lexical principles.

Using relevance theory we can account for the interpretation of (1) as follows. Consider (2):

- (2a) John is eating marbles.  
(2b) John is eating (=1).

Adopting Rizzi's proposal and modifying it slightly, let us assume that there is a grammar-internal lexically governed interpretative rule applicable to (2b) which spells out the implicit object of *eat* as an argument in the semantic representation (2b')



(2b') John is eating *x*.

Unlike Rizzi, let us assume that *x* is a variable and can be interpreted freely as, for example:

- (i) marbles
- (ii) food
- (iii) . . .

Potentially, (1/2b) has at least two readings with distinct truth-conditions. The principle of relevance predicts that the more relevant of the two readings will be selected. Relevance is a function of the contextual effect of the proposition expressed in the utterance and the processing cost. Under interpretation (i), *x* = marbles, and (1/2b) will therefore have the same contextual implications as (2a), for example:

- (3a) He must be starving.
- (3b) Stop him.
- (3c) The child-minder is careless.
- (3d) . . .

Under interpretation (ii), *x* = food, and (1/2b) will, for example, have the contextual implications (4):

- (4a) Do not call the child-minder now.
- (4b) His mother is busy.
- (4c) Do not distract him.
- (4d) . . .

Although both interpretations have some contextual effect, the contextual effect to be derived from the propositions in (3) will be larger than that obtained from the propositions in (4). This, however, does not mean that interpretation (i) (*x* = marbles) will be selected and that (ii) (*x* = food) will be discarded. On the contrary, in spite of its dramatic contextual effect, relevance theory predicts that the marbles interpretation (i) will be discarded and that the food reading (ii) will be selected. The argument will run roughly as follows: If the speaker had wanted to convey that John is eating marbles, he would not have been observing the principle of relevance in uttering (1/2b). Since *x* is a variable, (1/2b) is subject to several readings. To recover the full proposition, the reader has to assign an interpretation to *x*. To attain the food reading, all the hearer needs to do is to access the stereotypical encyclopedic information stored under the concept "eat," where he will find connections with the concepts "food," "meal," and so on. In order to obtain the marbles reading, the hearer has to create a novel connection. He cannot trade on the encyclopedic information directly available. Thus (i) becomes more costly to access.

Given first that the food interpretation of (1/2b) in a "normal" context will have a contextual effect (cf. (4)), there will be no need to access alternative and more costly interpretations. The food interpretation is the first interpretation in accordance with the principle of relevance.

Second, (2a) would be a manifestly more relevant stimulus to convey the "eat marbles" proposition. The specification of the object in (2a) saves the hearer the effort of disambiguation and directs him straightforwardly to establishing a (novel) link: "eat" – "marbles."

This account predicts that (1/2b) *John is eating* will "normally" be given the interpretation of "eating food" without there being any need to derive the interpretation on the basis of grammatical principles. Only in exceptional circumstances would the marbles reading be less costly than the food reading.

On the basis of my discussion of indefinite object drop, I hope I have shown that the emergence of a very precise and coherent theory of utterance interpretation means that one needs to reconsider the division of labor between grammatical principles of interpretation and the general principles of utterance interpretation located outside the grammar proper. In this way Sperber & Wilson's work is a challenge to syntacticians as well as pragmaticians.

## Relevance: Computation and coherence

Elizabeth Hinkelman

Computer Science Department, University of Rochester, Rochester, N.Y. 14627

Computational linguistics challenges students of language to provide precise and detailed theories but, unsurprisingly, few succeed. Though *Relevance* presents pragmatic phenomena in refreshing range and nuance and is concerned with computational issues, it solves no computational linguistic problems and describes a theory that suffers more from flaws and inconsistencies than from "abstractness."

Sperber & Wilson (S&W) want a pragmatics engine that is more general and plausible and less rigid than Grice's maxims. The term *ostension* (really Grice's *communicative act*) advertises that the concept is grounded in the individual's inherited or learned ability to recognize important stimuli, including intentions. But this pleasant promise is unfulfilled and may have to remain so, especially because intention recognition rather resembles general cognition. Likewise, *manifestness* claims to replace *belief* by grounding it in sensation and deduction, trying thus to eliminate the infamous regress of mutual belief. S&W implement mutual manifestness by propositions having some associated rule to identify who has access to them. This doesn't suffice: Persons with access must still infer or assume that others have the rule and use it, that the others observe these things about the rest, and so on for multiagent planning and discourse (Rule manifestness does not guarantee rule use). Mutual manifestness does not prevent regress: default formulations of mutual belief are more satisfactory (Perrault 1987).

Now consider *relevance*, S&W's technical term describing the usefulness of a proposition or perception as the effect it has on the strength of other propositions (informativeness) and the effort expended in computing the effect. The particular comparative algebra given for relevance has snags. Values may be compared just when either effort or effect is already equal (pp. 125, 132). (Formally, if *A* and *B* are totally ordered sets with *a* ∈ *A* and *b* ∈ *B* standing for levels of effort and effect respectively, a relevance value is the ordered pair *(a, b)*. *(a, b) > (c, d)* iff (*a* = *c*, *b* < *d*) OR (*b* = *d*, *a* > *c*.) There is no clause favoring values having both lower effort and larger effect (*a* < *c*, *b* > *d*), although it is clearly needed (see p. 168). These limits on comparison make each value incomparable with nearly all others. Hence acts of communication having multiple interpretations can easily have no most relevant interpretation, despite the claim of a unique solution (p. 168; also implicit in the Principle of Relevance, which refers to optimality).

The principal claim is that our brains are optimized to compute the most relevant implicatures first: Thus S&W's Deductive Device proceeds by making inferences in order of efficiency. But this is circular. The fact that an inference is the first one computed makes it relevant via efficiency, even though it may not be informative at all. S&W use efficiency (pp. 167–68) to impose a total order on interpretations found, as if this ordered them by relevance; but this *guarantees* that their relevances are incomparable (except among equally informative interpretations, when the Device stops with the most efficient one). The algebra of relevance cannot be patched by optimizing efficiency with respect to itself. One should optimize efficiency with respect to informativeness, and then rewrite the algebra to allow requisite comparisons. Note that optimality, informativeness, and relevance have an implicit teleology, here apparently conflating evolution, individual survival, the individual's long- and short-term goals, and impulses.

Another problem appears when we consider computing implications in context. My attempt to diagram S&W's "system" (Figure 1), shows the Deductive Device taking as input an utterance and a context, determined by the context selector. The context selector chooses that context yielding optimal relevance, in an unspecified manner: rounded boxes are mod-

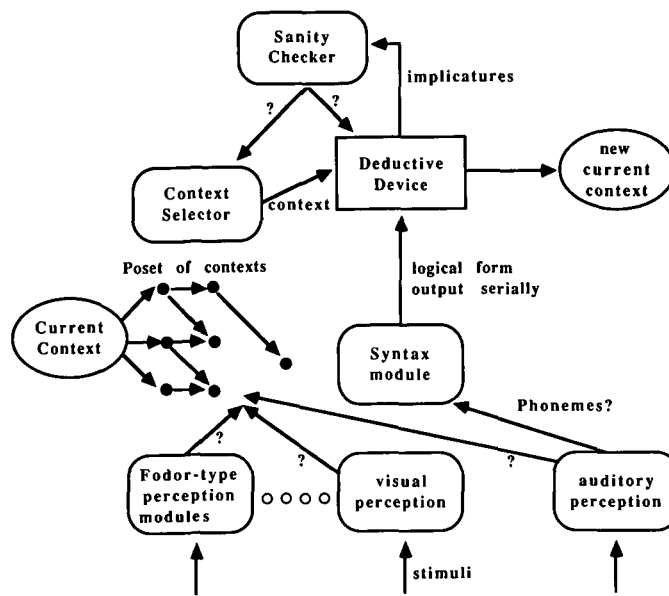


Figure 1 (Hinkelman). A relevance computer.

ules without specified workings. Since it has no access to any strength, manifestness, or relevance values (they are “non-representational”), nor to the utterance, it has almost no basis for decisions.

S&W might make some progress here by trying to be clear about the different levels of representation in their system. There seem to be three: a “hardware” level, the inference level on top of it, and the level of consciousness. Some “nonrepresentational” quantities are available as chemical or electrical quantities at the hardware level and are thus able to guide inference without ever reaching consciousness. (I suggest that noncomparativeness is an artifact of conscious introspection about the hardware level.) If strength appeared chemically at the hardware level, then context selection might be understood simply as the existence at the inference level of a pattern of hardware electrical activations: The partial order of augmentations could be activated by the inference process. S&W could profitably look at connectionist models as a starting point for specifying their demands on the hardware level (Waltz 1987). The inference level could be compared against such work as Frisch (1985), which describes an inferential memory model. It could also be improved by detailing the contents (or at least the interfaces) of the rounded boxes in my diagram. When S&W sort their desired properties by level and specify the interfaces, they will improve their model immensely. Sometimes circular arguments are put right in this way.

Most of the unconscious inference operations are performed in S&W’s model by the Deductive Device. It uses only elimination rules to guarantee that from a finite set of propositions (and a finite set of rules) it will eventually perform all possible inferences and stop. However, it can still fail to terminate: Any logic with *modus ponens* and function terms can regress infinitely (consider: She is a citizen if her mother was a citizen). Fortunately, if processing time is bounded, the worst that will happen is that no interpretation will be found. The Deductive Device updates the strengths (not probabilities or truth values) of propositions nonmonotonically, although strengths are “not represented.” The relationships among strength, manifestness, and relevance remain unclear.

Some general considerations that motivate alternative systems are not discussed in *Relevance*. The interplay of lexical, syntactic, and pragmatic processing prompted Charniak’s so-called neat theory of marker-passing (Charniak 1986). The massive fine-grained parallelism of the brain motivates connectionist models (Cottrell 1985; Dell 1986), as does the im-

plausibility of passing around entire parse trees. Anaphora, reference, ambiguity, and error correction are addressed by the plan-based approach (Grosz & Sidner 1986; Allen 1987). The *Relevance* model may also fail to generalize readily to language generation or general cognition as claimed.

*Relevance*, then, serves computational linguistics best as a reminder that the range and subtlety of pragmatic phenomena are not yet covered smoothly by one theory; we must consider the slippery substrate of human intelligence when modeling its performance. S&W’s theory does not improve on Grice for consistency or technical detail.

## Grammars as input systems

Ruth Kempson

School of Oriental and African Studies, University of London, London WC1E 7HP, England

It may surprise many that a theory of communication and cognition can have consequences for the logically prior concern of articulating the language specific knowledge a speaker brings to general cognitive processes. Yet Relevance Theory (RT) imposes precise requirements on the input to the process of utterance interpretation provided by the grammar. Indeed, the hallmark of RT is the wealth of specific predictions that emerge as consequences of its central tenets, despite their own generality. I give four consequences:

1. The concept of content articulated in the grammar is representational (being a representation in the language of thought (Fodor 1983) – hereafter LT). This claim is compatible with Chomsky’s (1981) theory of Universal Grammar (UG), but in conflict with all truth-theoretic theories of content (cf. Dowty et al. 1981 and references cited there).

2. The logical-form representations associated with sentences of a language as part of its grammar are not syntactic constructs of the same system as the natural language itself but are syntactic constructs of LT (as a consequence of viewing grammars as input systems).

3. Truth-theoretic content is assigned to propositional forms, expressions in LT completed by a pragmatic process, and not to natural language expressions directly (this a consequence of logical forms underdetermining their propositional content).

4. Any aspect of interpretation that is pragmatic will involve

structural representations as values and will be subject to the principle of relevance – hence likely to have values which vary across limits set by the requirement of minimal processing cost.

Each of these consequences is strongly confirmed. It is well known that the ascription of truth-theoretic content to sentences is not compositional as defined over linguistic expressions, but involves a pervasive dependence on contextual factors (Partee 1984). If we assume that the articulation of such content is not part of the speaker's language faculty but is provided only for LT propositions, then the dependence of truth-theoretic content on contextual parameters is predicted. (Ascribing content to expressions of LT is no less a problem, but it is one that has to be faced anyway.)

Prediction (4) is confirmed by such anaphoric processes as using pronouns. These display a heterogeneity which has to be stipulated on truth-theoretic assumptions of linguistic content, their values ranging over any represented information accessible at minimal processing cost to the hearer. On the assumption that truth-theoretic interpretation is fixed pragmatically, constrained by the principle of relevance, we can provide a unitary account of the relatively weak linguistic content intrinsic to anaphoric expressions and we can directly predict the range of truth-theoretic values they display (Kempson 1986).

Predictions (2) and (1) jointly provide a rich vein of auxiliary predictions and confirmation. Recent studies in phonology, morphology, and syntax have provided increasing evidence that if a unitary explanation of the interpretation of natural language structures is to be given, it must assume at least one level of structure distinct from PF structure, s-structure, and LF (as defined in government/binding theory; Chomsky 1981). Gussenhoven (1983) has provided evidence that, though stress assignment interacts with hierarchical structure to determine interpretation, the level of structure required is some "semantic" one involving concepts such as predicate, argument, and focus domain, rather than that provided by syntax. In morphology there are many cases across widely varying languages in which the morphological structure does not provide the required basis for interpretation of the structure (Pesetsky 1983; Sadock 1985; Williams 1981). A simple example is *generative grammarian*, which requires the compounding of *generative* and *grammar* prior to interpreting the affix *-ian*, thus assuming for purposes of interpretation the structure: [[*generative grammar*] *-ian*].

And there are the problematic "Reconstruction" phenomena in which the Binding Principles (Chomsky 1981) constraining the interpretation of reflexives and pronouns apply not to LF structures but to the output of an "interpretation" process which in effect reverses Move WH (Williams 1986) – exemplified by *Which of each other's pictures do you think that Bill and Mary deliberately ruined?* None of these problems has been seen as related to the others, and the solutions proposed postulate various arrays of double-structure assignment. Yet each of the phenomena is predicted from Relevance assumptions, and for the same reason. In order for the claim that grammars are input systems to a central LT to be nontrivial, the mapping from phonological/morphological/syntactic structure onto logical form must carry the natural language structure onto a discrete structure in at least some cases. The cases which have emerged in these different areas are those in which the LT structure specified as the logical form of the expression is not identical to that provided by the natural language structure. The prediction of their existence is thus an immediate consequence of viewing grammars as input systems to a central device with its own LT (Kempson, forthcoming b). Furthermore, RT predicts the property of Reconstruction (embarrassing to GB) of not being subject to well-formedness conditions imposed on Move  $\alpha$ . A process-mapping natural-language structures nontrivially onto structures of a discrete LT is by definition not restricted to the well-formedness conditions associated with processes internal to natural language.

The explanatory power of RT also extends to the functional motivation for properties of UG. Properties of UG such as the principle of Full Interpretation (that no expression at LF is contentless – Chomsky 1986) and the Theta criterion (providing a unique association between syntactic and logical arguments) can be seen as motivated by the constraint of maximizing relevance (Kempson, forthcoming a). This is not to deny their status as properties of UG, but rather to provide an explanation of their existence, the language faculty being hardwired to provide analyses of linguistic strings with no wasted processing effort.

Thus RT provides new and precise solutions to a whole range of old problems. In each case these solutions are not piecemeal devices but a coherent part of an overall explanation. There is now the need for a new linguistic theory articulating the mapping from phonological/morphological forms via syntactic structure onto LT expressions in order to give these putative solutions detailed empirical substance. But the general message is clear enough: RT opens up new avenues of research not merely in its own domains of philosophy of language, pragmatics, and psychology, but in semantics, syntax, and even morphology and phonology. In so doing, it proves itself to be a theory which offers considerable intellectual rewards.

## Implicature explicated?

Stephen C. Levinson

*Department of Linguistics, University of Cambridge, Cambridge CB3 9DA, England; Department of Linguistics, Stanford University, Stanford, Calif. 94305*

Sperber & Wilson (S&W) claim that a single principle of relevance (R) can account for all pragmatic inference; but they fail to define R clearly enough to make this an empirical claim. R is a function of (measuring) contextual effects (E) balanced by processing costs (C) – related, say, as  $R = E/C$  (though not thus subjectively assessed).<sup>1</sup> This allegedly yields determinate interpretations. Yet how the equation is solved varies: Sometimes R has a predetermined value (V), requiring contexts to be expanded until V is satisfied (p. 142); sometimes R is a comparative measure, selecting the best interpretation (pp. 144, 153); sometimes C has a threshold value such that the first accessible E-yielding context is automatically selected (p. 178). But if only R is fixed there is no determinate solution; if C is fixed too, then comparing interpretations may exhaust C without producing any solution. (A puzzle: If  $C = n$  and  $E = m$  for interpretation 1, but  $C = n + 1$  and  $E = m + 1$  for interpretation 2, then the cost of each will be the cost of comparing them, viz.  $2n + 1$ , leading to the conclusion that one should always pick maximum E regardless of C.)<sup>2</sup> Finally, without an objective measure of C, how can R be empirically tested?

There are further unclaritys: Is R a measure applied to surface structures (p. 204), semantic representations (SRs), or explicated logical forms (LFs) – or even recursively to LFs plus implicatures, or to all of these? If the answer is to more than one of these, then the claim that a speaker should optimize R makes no clear prediction about how the speaker should distribute "meaning" between the "encoded" and "implicated," and thus how a recipient can best interpret the utterance.

Turning to implicature: S&W's prototype cases are those corresponding to Grice's (1975a) relevance implicatures. S&W admit that their cases mostly involve implicated premises; they also assert (p. 83) that deduction is the key to understanding nondemonstrative inference, including implicature. But because implicated premises cannot be deduced, the very prototype relevance-implicatures cannot be accounted for. Nor will "accessibility" provide heuristics (see p. 201) in cases in which the implicated premise is, as often happens, not a recipient's (e.g. A's) assumption (cf. p. 194):



- A: "Would you drive a Zordia?"  
 B: "I wouldn't drive ANY expensive car."  
 Implicated premise: "A Zordia is an expensive car."

This failure is played down by introducing implicated premises as if they were contextual assumptions (e.g., counterfactual premise [5]e, p. 122). "Real" (pretheoretical) relevance is largely about guessing the speaker's goals (see Allen 1983; Dascal 1977; Holdcroft 1987), not maximizing information for R.

Turning to explicatures: These include all those generalized conversational implicatures (GCIs) whose understanding is the signal achievement of neo-Gricean pragmatics. GCIs are default pragmatic inferences which may be canceled by specific assumptions, but otherwise go through. S&W hardly engage this work (except on p. 262, note 4); yet they imply that R automatically accounts for GCIs as explicatures. The onus is on them of course. S&W consider that the neo-Gricean theory of GCIs (Atlas & Levinson 1981; Gazdar 1979; Horn 1972, 1984; Levinson 1983) is a result of brainwashing by the "code-model" of communication (pp. 36ff). But they make an important mistake in equating default but defeasible inferences with automatic decoding. Indeed, any theory of inference that takes processing cost into account must surely employ default inferencing – witness its almost universal use in computer models of language use.

The theory of GCIs makes clear predictions. If R also made clear predictions, they would almost certainly be different. The following may be a case:

- A: "If it's possible that the spy has *more* than two passports, he may yet escape."  
 B: "He has two passports."

GCI theory holds that *two* means "at least 2" (allowing the possibility of more) but also has GCI "at most 2"; the theory thus predicts that B implicates "He has only two passports," thus suggesting that A's fears are groundless, and (perhaps) that the spy will not escape. If "two" means "at least two," R theory would presumably predict that there would be no explicature "at most two" because that would rule out the low-cost contextual implication "he may escape," thereby lowering R. Thus B's utterance should implicate "He may escape." My intuitions favour the GCI predictions.

S&W may complain that this argument allows implicatures to effect explicatures, whereas in fact explicatures form a prior independent level (see Kempson & Cormack 1981 and the effective rebuttal in Horn 1985). The explicature/implicature distinction is salutary as a means of emphasizing the pragmatic determination of LF, but it is not clear. The only criterion offered is that explicatures must contain the encoded SR or LF as a proper subpart (p. 181); but many implicatures meet that condition:

- A: "If Thatcher has won the election, she'll have won three times."  
 B: "Thatcher has won."  
 Implicature: "Thatcher has won three times."

Nor can explicature-calculation always precede implicature-calculation (cf. p. 179); for example, disambiguation must be guided by contextual effects.

S&W are to be congratulated for emphasizing the role of pragmatic inference in communication in general, and in propositional determination in particular. But other pragmaticists are unlikely to agree with monadic R theory (Sadock 1986). R is an informational measure; but Atlas and Levinson (1981), Horn (1984), and others have argued for the necessity of two counterbalancing informational measures. And, as noted, pretheoretical relevance is largely about satisfying the goals of others, whereas Quality (pace S&W) remains unreduced. For R to encompass all these other principles is to stretch it too thin.

Meanwhile, to compare R to, say, GCI theory, the newer formulation needs much clarification: Since R is a processing theory, how about a computer implementation?<sup>3</sup>

NOTES

1. This paragraph is not meant just to recycle the critique by Gazdar and Good (1982) (although to my mind S&W still haven't dealt fully with that), but rather to ask for clarification about how R can be computed.
2. S&W (p. 131) suggest that heuristics will allow estimation without computation. We need detailed hypotheses, however. Without them the puzzle looms large; with them, predictions might be made about specific kinds of misinterpretation.
3. One attempt at computer implementation is being made by R. Poznanski under K. Spark Jones of the Computer Laboratory, University of Cambridge.

The multidimensionality of pragmatics

James D. McCawley

Department of Linguistics, University of Chicago, Chicago, Ill. 60637

Sperber & Wilson (S&W) fairly quickly dismiss the notion of "mutual knowledge," for reasons that relate to problems created by the specific conception of mutual knowledge that they consider: one which is defined in terms of ordinary knowledge and thus requires reference to propositions of the form "A knows that B knows that A knows that B knows that . . ." ad infinitum.

Suppose, however, that one takes mutual knowledge to be something really mutual – that is, that one speaks of knowledge on the part of a group of two or more persons as a whole and does not attempt to reduce such knowledge to knowledge on the part of each of those persons individually. Mutual knowledge, thus understood, is not knowledge that the participants in a conversation hold individually about one another but rather knowledge they possess jointly. The identification of propositions as mutual knowledge does not, then, require the "infinite series of checks" that S&W speak of (p. 18). The following are unproblematic sources of mutual knowledge: (i) propositions that a participant has asserted during the conversation, unless a proposition has been either retracted by the speaker or challenged by another participant; (ii) information that is "manifest" (in the sense of S&W, p. 39) to the participants jointly, including information about what has been said so far in the conversation (e.g., the proposition that John said that pigs have wings is mutual knowledge, even though the proposition that pigs have wings isn't, if Mary challenged John's statement to that effect), and (iii) information commonly regarded as manifest to the members of a group in which the participants share membership. This last source of mutual knowledge is discussed insightfully by Prince (1978), who notes as an example that the identity of the mayor of Philadelphia is mutual knowledge if the participants in a conversation can reasonably assume that they all are residents of Philadelphia, as when strangers strike up a conversation on a Philadelphia bus. This last source of mutual knowledge is what licenses the uttering of (1) even if the speaker does not know whether the addressee has read *Sense and Sensibility*:

- (1) It's the sort of scene that would have made Marianne Dashwood swoon (*Relevance*, p. 44).

*Sense and Sensibility* has the status of a literary classic; and in circles in which familiarity with the classics is presumed, the identity and personality of their major characters is "manifest" and thus will count as mutual knowledge, in the sense that a person is entitled to utter (1) without first having to say who Marianne Dashwood is. Note that on this understanding, a proposition can be mutual knowledge in a given conversation even if not all of the participants know the proposition and even if the proposition is false: Many false propositions are manifest.

Despite S&W's dismissal of mutual knowledge, suggesting

that they believe their approach to offer a viable alternative to analyses that have recourse to the notion of mutual knowledge, *Relevance* is surprisingly lacking in discussion of phenomena that other authors have dealt with in terms of mutual knowledge, such as pragmatic presupposition in the overlapping accounts given in Karttunen (1976) and McCawley (1979; 1981), which interpret a pragmatic presupposition as a requirement that the presupposed proposition be mutual knowledge.<sup>1</sup> One particularly important area of pragmatics in which I consider mutual knowledge an indispensable part of any analysis is the assignment of referents to definite noun phrases. As Donnellan (1966) has pointed out, a sentence containing an expression *the X* need not be given an interpretation in which the referent of the *X* is an *X* or even is believed by the speaker and the addressee to be an *X*. For example, *the man drinking the martini* may identify a person who is drinking water even if another person is in fact drinking a martini. Similarly, in a conversation between Daddy and Johnnie, the expression *the Easter Bunny* can recur in the speech of both even if neither believes that there is an Easter Bunny or that the other believes there is an Easter Bunny. What licenses the use of such noun phrases is that the mutual knowledge for the particular conversation includes the proposition that the given person is drinking a martini or the proposition that there is an Easter Bunny.<sup>2</sup> I in fact suspect that there is little overlap between the class of problems that can plausibly be dealt with in terms of mutual knowledge and the class that can plausibly be dealt with in terms of S&W's relevance; my complaint with S&W is not that they have failed to solve problems they should be expected to solve but that they falsely claim to have discredited an approach to which theirs is not even clearly an alternative.

Turning, finally, to S&W's notion of relevance (which should be judged on its own merits and not as an alternative to mutual knowledge), I remain unconvinced either that it is a unitary notion (at most, in a technical sense have S&W reduced Grice's maxims<sup>3</sup> to one) or that their concept corresponds very closely to intuitive notions of "relevance." Richness of inference and complexity of processing remain independent dimensions, and S&W do not even give any clear examples of a trade-off between the two. The first of these two dimensions is a plausible first approximation to the intuitive notion of "relevance," in that one's most effective response to an objection that something is not relevant is usually to cite inferences that it enables one to draw. However, such objections are never about relevance *simpliciter*, whatever that might be, but relevance *to* something, generally to either a question or a subject area. Propositions that add to the inferences one can draw can easily lead one off on tangents (*ad hominem* arguments often license far more inferences about their victim than about the point at issue). S&W's brief discussion of the notion of "topic relevance" (pp. 216–17) dodges the issue by merely stating that "Topic-relevant utterances are only a subset of relevant utterances" and that "It is the notion of topic relevance which is derivative." The former statement may well be correct (I give S&W the benefit of the doubt about the fuzzy set theory that would need to be invoked to reconcile "subset" with their notion of relevance) but is beside the point, since it is the larger set whose significance is more in need of justification; and their second statement is dubious in view of the weakness of their attempt to derive "topic relevance" from "relevance": "relevant (in our sense) in a homogeneous context derivable from a single encyclopaedic entry" p. 216; a single encyclopaedic entry isn't my idea of what defines a "topic."

S&W's use of the word *irrelevant* is often at sharp odds with its use in ordinary language, as when they describe blatantly false propositions as irrelevant<sup>4</sup> (p. 120) or when they speak of an utterance expressing a proposition already known by the addressee as "irrelevant to him" (p. 160); "Nice day, isn't it?" comes out as "irrelevant" in S&W's scheme. I am intrigued by S&W's highly original idea that relevance is a matter of "new"

inferences that a proposition allows one to draw by means only of "elimination" rules. However, I am disappointed that that genuinely appealing idea gets lost fairly soon in their attempt (misguided, in my opinion) to project onto a single dimension the diverse factors that influence how persons interpret the utterances that they hear and see.

#### NOTES

1. This is an oversimplification: Sentences are not interpreted relative to the mutual knowledge *per se* but relative to mutual knowledge supplemented by other propositions supplied by the linguistic context. Thus in understanding the sentence "If Reagan appointed Jesse Jackson Secretary of State, he would regret that he had named a civil rights leader to the Cabinet," one evaluates the main clause relative to the mutual knowledge supplemented by the proposition expressed by the *if*-clause (i.e., the protasis of a conditional sentence provides "supplementary context" for the apodosis).

2. For arguments against Russell's analysis of definite descriptions and in favor of one that relies on mutual knowledge, see Lewis (1979) and McCawley (1985).

3. A better comparison would be with Horn's (1984) revision of Grice. Horn's two maxims are direct counterparts of the two parts of S&W's single maxim (*Relevance*, p. 125).

4. Truth is a third dimension that S&W conflate under "relevance." I doubt that there is any tradeoff between this dimension and the other two. For example, in Bertrand Russell's (1905) ambiguous example, "I thought your yacht was longer than it is," it is presumably easier to compute the reading with the embedded contradiction than the reading in which the *than*-phrase has wide scope; but ease of processing does not lead one to settle on the obviously false reading.

### Logical competence

John Macnamara

Department of Psychology, McGill University, Montreal, Canada H3A 1B1

Sperber & Wilson (S&W) succeed in conveying an impression of the mind as an active processor of information designed to serve practical aims. This correctly sets the mind apart from a logical system that presents itself as already objective and complete before logical inquiry begins. The problem is to be true to both the practical purposes of information processing and to logical intuition. I fear, however, that S&W are not true to logical intuition. My remarks will be confined to Chapter 2, which discusses "inference."

S&W have a somewhat odd idea that a concept in memory stores three types of information: logical, encyclopedic, and lexical (p. 86). They then go on to say "that the 'meaning' of a word is provided by the associated concept" (p. 90). Their idea seems to be that the meaning of *giraffe*, say, consists of (a) a set of deduction rules in which *giraffe* can figure, (b) a set of beliefs such as that giraffes have long necks, and (c) "information about its syntactic category membership and co-occurrence possibilities, phonological structure, and so on" (p. 90).

S&W are skimpy on what "cooccurrence possibilities" are and how they might be represented. *Giraffes live on the moon* is a perfectly good sentence. Does it violate the cooccurrence rules for *giraffe*? If not, those rules would seem to have little to do with meaning, which surely has an intimate connection with the truth conditions of sentences in which a word occurs. If the sentence does violate the cooccurrence rules, we need to be told very much more than S&W tell us about the nature of those rules, including how they are learned and stored. The key element in the meaning of the word *giraffe* has not really been addressed.

Matters are worse. Inessential properties of giraffes have been bundled into the meaning of *giraffe* through the encyclopedic information; they thus become analytically true of the creatures. For example, the relevant encyclopedic entry is sure to contain the information that giraffes live in Africa. Yet it

is not analytically true of giraffes that they live in Africa, as it is analytically true of spinsters that they are unmarried. Things are getting seriously muddled. I tried to straighten some of them out in Macnamara (1982, Chap. 12).

S&W's great claim about the deductive component of a concept is that it involves only elimination rules, no introduction rules. Here is an example of an introduction rule: Given *John is tall* and *Mary is small*, the rule for *and*-introduction permits the inference of *John is tall and Mary is small*. Given the latter sentence the rules of *and*-elimination permit the inference of *John is tall*. S&W, it should be noted, are not limiting their claim to the logical component in a concept. They state quite generally, "Our hypothesis is that the human deductive device has access only to elimination rules" (p. 97).

This is profoundly puzzling. How could there be anything for elimination rules to work on if there were no introduction rules? There would be no sentences with *and* in them unless there were some sort of *and*-introduction rule. So what might such a rule be? Well, it might be a syntactic rule that says that if *p* and *q* are well-formed sentences, so is *p and q*. The rule would need to be supplemented by a semantic rule that specified the conditions under which *p and q* is true. But together the syntactic and semantic rules are tantamount to an *and*-introduction rule. What seems to have been worrying S&W was the problem of shutting off the "human deductive device." The source of the problem is that, granted the truth of *p*, conditions are satisfied for deducing (*p & p*) and (*p & (p & p)*) and so on ad infinitum – a series of pointless deductions. But the problem does not arise in ordinary discourse with elimination rules: They quickly leave themselves with nothing to eliminate. S&W's solution of eliminating *and*-introduction, however, either fails because it merely shifts the dangerous generative capacity to the syntax or succeeds by brute force. If there are no *and*-introduction rules of any sort, there is no danger of a pointless string of conjuncts; but neither is there anything for the *and*-elimination rule to work on. S&W seem to be solving the problem in the steering by taking out the spark plugs.

I believe, however, that the trouble is deeper – that what we have seen are merely a few symptoms of a serious malady. S&W proceed as though they have no obligation to account for the existence of logic of the type one finds in the textbooks of formal logic, as though formal logic bears no intimate relation to the logic of everyday reasoning. Yet the only access a formal logician has to the truths of logic is through those mental processes that yield logical intuition. Formal logicians base themselves on some of those intuitions, those they find most suitable for their purposes. Everyday reasoning embraces everything that the formal logician works with and much more besides: presuppositions, Gricean implicatures, ambiguity, vagueness, and so forth. So an account of everyday reasoning must at least give the psychological grounding of all those basic intuitions upon which formal logic rests.

In Macnamara (1986) I attempt to work out the implications of what I have just said. I suggest that we take formal logic as providing us with a (limited) statement of logical competence, leaving room for logical error while being true to basic logical intuitions. The trouble with formal logic is that it derives in modern times from the study of arithmetic and thus treats a domain of eternal and unchanging individuals. The number three does not get fat or die. Most discourse, however, deals with a domain of changing and transitory individuals. More than anything else what psychology needs is a logic of such transitory individuals. Though the subject of relevance is not a psychologically trivial one, S&W do not, I fear, provide the requisite logic in *Relevance*.

## What Peter thinks when he hears Mary speak

Ruth Garrett Millikan

Philosophy Department, University of Connecticut, Storrs, Conn. 06250

*Relevance* is a book very densely packed with ideas. Many of them are novel and a good number are, in my opinion, correct. But I find others highly puzzling. I offer a couple of examples here. I intend them not so much in the spirit of criticism as in the hope that Sperber & Wilson (S&W) will unpuzzle me – with a bit of private tutoring. Thought number three, however, I urge with more confidence: It is backed by Millikan (1984a, Chap. 3).

First, according to S&W, Peter (the hearer) determines Mary's (the speaker's) meaning not by reference to some subset of background assumptions mutually manifest to him and to Mary, but by reference to *all* assumptions manifest to him in their current degrees of manifestness. He does this by finding the interpretation of Mary's speech that is most relevant to him. Does Mary not have to anticipate what will be relevant to Peter then? Yet what is relevant to Peter depends on what Peter knows and on how his knowledge is organized. Possibly the organization of Mary's mind is like that of Peter's, but will she not also have to know everything that Peter knows and know that he knows it? Also, do S&W intend to suggest that Peter need not think at all about Mary's knowledge? Suppose that Mary refers to the woman Peter was just with as "your girlfriend." In truth it was Peter's sister; moreover Peter does have a girlfriend. Will Peter not better understand Mary by knowing or inferring what Mary does and does not know?

Second, as S&W note, in most cases we tell what people's intentions are by observing the effects of their behaviors. Yet a person's behavior always has multiple effects, most of which are not intended. For example, as I move to look in the refrigerator it may be that, unintentionally, I advance three feet toward London, make the floor squeak, and whistle Eisenhower's favorite tune on John Adams's birthday. To isolate people's intentional actions from among all the irrelevant things they do I must first have an opinion about what it might be *sensible* for them to do and how it might be sensible for them to go about doing it. Now as S&W describe communication, Mary gets her informative intention recognized by knowing how to manipulate Peter's psyche, how to produce certain effects in it. She knows what stimulus will produce a roughly determinate sort of thought chain, one link of which will be a thought of the proposition (*p*) she wants Peter to believe. To communicate that *p*, Mary must first know how to make Peter think of *p*. Peter then recognizes *p*, from among all the other things Mary's stimulus also made him think of, by the fact that it is the first thought in the chain that is relevant enough. So it seems that Peter recognizes Mary's intention in the ordinary way, except that he makes this assumption: The only effects it would be *sensible* for Mary to try to produce on his beliefs are ones highly relevant to him – relevant in S&W's technical sense of "relevant" (and, of course, the only sensible way for Mary to go about this is with as much processing efficiency as possible).

But why would Peter assume this? Certainly it is true that Peter will not be in the habit of paying much attention to Mary's remarks unless it often pays him to process them and pays more than the processing cost – and that Mary will do well to make enough of her remarks pay enough to keep Peter betting. But are there not other ways for Mary to pay Peter than by telling him something (technically) "relevant"? Like telling him something that pleases him or amuses him or is intrinsically interesting to him, say – assuming that he loves Mary – telling him how she feels today or that she loves him? (If these things are, by some argument, *automatically* the most "relevant" in S&W's sense, then I need to have it spelled out for me why.) I also want to ask, might not Mary lapse occasionally and say something to Peter for her own sake only, for example, "Please pass the salt"? (Are another's requests automatically most "relevant"? Why?)



In brief, why won't just the ordinary way of interpreting people's intentions do for language interpretation too? Won't it be enough for Peter to recognize Mary's informative intention if he uncovers a proposition that it would be sensible for her to want to impart in that context and to impart in that way? Compare: Having asked for the salt, Peter readily interprets Mary's outstretching arm as manifesting an intention to pass the salt; having asked for some information, "Do you want coffee?" Peter readily interprets "It would keep me awake" as manifesting an intention to supply the information he requested – no, she wouldn't like coffee.

Third, and finally, if the contemporary position is correct that belief-desire talk expresses a folk theory of the mind instead of expressing concepts and knowledge that originate as introspective epistemological "givens," is it really plausible that extremely young children represent to themselves that the speaker *intends* that they should *believe* that *p* in the process of understanding their very first sentences? Notice that it won't help to claim that very young children do not communicate in a fully adult way. For it is not just the act of communicating that requires Peter to think of Mary's intentions. According to S&W (and they have lots of good company), the only way for Peter to become *informed* via Mary's speech is, in the usual case, by making an inference of this kind: Mary intends me to believe that *p*, and I trust her, so *p* must be true. It seems unlikely that tiny children have the requisite concepts to make this sort of inference. But if it is possible for very young children to become informed by speech without using belief-desire theory, should it not be equally possible for adults? Perhaps there are some kinds of sentences that, at least when used in some kinds of ways, can be interpreted by Peter without his having to go through inferences about Mary's intentions?

Indeed, S&W's theory of how Peter determines the "explicitures" of Mary's utterances seems perfectly tailored to explain how Mary's informative intentions might get fulfilled without Peter's thinking of them. Reference assignment and enrichment of propositional content are both achieved without reference to Mary's intentions. In most cases, all Peter needs to do to fulfill Mary's informative intention is to believe whatever Mary's utterance makes him think of first, and to move on from there by inferences natural to him. Of course, it has been traditional to hold that no adult would be so gullible as just to believe what he was told without going through an inference that makes explicit his evidence, including his trust in the speaker. But S&W are not afraid to smash idols. We have all been deeply indoctrinated into foundationalist epistemology, but *psychologically* speaking, does it not seem likely that mistrust of a speaker inhibits the mechanisms of normal belief transfer rather than that trust in a speaker is used as a premise for each act of belief fixation that results from interpreting speech? If Peter sometimes understands without believing, perhaps this is merely the interruption of a normal process of belief fixation before completion, not a separate act of belief fixation – that is, fixing the belief that Mary intends him to believe that *p*. (For a discussion of how to relieve Mary also of the onerous task of having to use belief-desire theory every time she speaks, see Millikan [1984a].)

Incidental note: Regarding another way to use the (very sensible, I think) notion that strength of assumption is functional and determined by the psychological origins of assumptions, see Millikan (1984b).

## On the search for relevance

Jerry L. Morgan and Georgia M. Green

Department of Linguistics, University of Illinois, Urbana, Ill. 61801

This is a meaty book that provides much food for thought for those people Sperber & Wilson (S&W) refer to as "pragmatists" ["pragmaticians" in the précis]. But settling the debates it is

bound to stimulate will not be a simple matter, because what S&W present is not really a theory of anything, but, in their own words, "a new approach to the study of human communication." Furthermore, their presentation is full of crucial uses of unexplained concepts such as "attention," "information," "representation," and many others. Consequently, they are not able to offer anything like the kind of rigor and explicitness that one is used to seeing in the best work in formal syntax or semantics. But precision is just what makes it possible occasionally to subject a theory to empirical test. The weakness is hardly unique to this book; it is probably inevitable, given the present primitive state of this kind of study, indeed of cognitive psychology in general.

It is to cognitive psychology (or maybe to what cognitive psychology may be some day) that S&W look for the conceptual bricks upon which their approach is built. We are fully sympathetic to this orientation, in that we think that the supposed convergence of cognitive and philosophical approaches to language understanding has been largely illusory, and that theories constructed using criteria of adequacy that appeal to logicians are likely to turn out in the long run to be irrelevant for empirical cognitive purposes. We say this in full awareness of the fact that philosophers are responsible for most of the fundamental insights into meaning and communication on which both contemporary philosophical and empirical studies are based. The treatment of deictic ("indexical") expressions by relativizing reference and truth to sets of primitive "indices" comes to mind as a good case of a theory that may well be logically impeccable, but cognitively useless. Another case is the common assumption, criticized by S&W, that there must be a single unified theory of communication into which human linguistic communication must fit. Still another is the occasional attempt to construe "context" in terms of the objective physical setting of an utterance, rather than in terms of cognitive states of the participants.

In sum, *Relevance* is a book with which we find it easy to agree on a number of important points. Nonetheless, we have some criticisms that bear on the notion of relevance that is the core of the work. Since what is presented is more "approach" than theory, we cannot be certain whether our objections result from crucial flaws in S&W's approach or from vagueness in the details.

To begin with, we have reservations about S&W's definition of "relevance." They treat the concept in terms of "contextual effects" having to do with the information conveyed by an utterance and with the relation of that information to information the hearer already has. (Space limitations prevent a fuller explication here.) For our purposes, the critical aspects of their characterization are these: It is *information* that has or lacks relevance, not utterances; and relevance has to do with (grossly simplified) the newness of the information, not with any non-linguistic usefulness the information may have for the speaker or hearer. But there is another way to look at relevance: as a property of information (or of utterances that convey information) that *matters*, that is *useful* for some nonlinguistic purpose of the hearer or speaker. If in the middle of a lecture on syntactic theory the lecturer says (truthfully), "I am allergic to chocolate," as far as we can see that information should count as relevant for S&W, but would be irrelevant on the usefulness account, since uttering it achieves no advance toward any current goal of the interchange. Now as S&W point out, their goal is not to elucidate commonsense usage of the word *relevance*, but to set out a concept that has some scientific utility, whether or not it matches the commonsense notion. But at the same time they explicitly appeal to evidence from intuitions of relevance. Our intuition is that, in test cases like the one above, the usefulness approach fits relevance intuitions better than S&W's approach, which as we understand it is closer to a characterization of "informative" than of "relevant."

S&W's view of the nature of ostensive communication similarly involves stretching the usual understanding of communica-

tion in such a way that its implications conflict with experience. S&W view ostensive communication as communicating “its own optimal relevance” – that is, as communicating

1. that the assumptions which the communicator intends to be made manifest to the addressee are relevant enough to make processing worthwhile (in that they are effectively informative – in their words, “the contextual effects achieved when . . . optimally processed are large” [p. 153] – and the effort required to process them is small) and

2. that the “ostensive stimulus” (the linguistic expression or gesture) is the most relevant one the communicator could have used to accomplish those effects.

This is a very Panglossian view in that it treats the fact that a speaker has just uttered something communicatively as presupposing that it is the most relevant utterance he could have uttered – everyone does the best he could possibly do in this best of all possible worlds. At the same time, in taking pains to contrast their view with Grice’s,<sup>1</sup> in which communicative behavior is treated as rational and as governed and interpreted by the same principles as other rational behavior (Grice 1975a, p. 47), Sperber and Wilson treat communication as “an essentially egotistic enterprise” wherein “the speaker tries to have the maximum possible effect on the hearer’s set of initial assumptions” (Wilson & Sperber 1981, p. 175). “Communicators and audience need no more know the principle of relevance to communicate,” they say, “than they need to know the principles of genetics to reproduce,” (1986a, p. 162) – not mentioning that speaker and interpreter have to “know” the Presumption of Optimal Relevance in order for their claims about how communication is accomplished (pp. 165–68) to be correct. The Presumption of Optimal Relevance appears to us to entail most of the maxims that follow from Grice’s Cooperative Principle, subsuming Quantity and Manner as well as Relevance, rather than replacing them with an altogether different principle, as S&W seem to suggest.

Furthermore, the claims embodied in S&W’s principles and definitions seem to make false predictions about certain kinds of implicatures common in ordinary discourse. It follows from S&W’s Principle of Relevance and their definition of relevance that “of all the interpretations of the stimulus that confirm the presumption, it is the first interpretation to occur to the addressee that is the one the communicator intended to convey” (pp. 168–69). (If it’s not, then the communicator failed to choose the most relevant stimulus – one which would have been interpreted right the first time. It is only the hedges *might* and *expected* in the explanation of what is “consistent with the principle of relevance” (pp. 166–67)<sup>2</sup> that keep this from being circular.

In any case, it seems to follow that ambiguities should never be perceived, and consequently it could never be rational behavior consistent with the Presumption of Optimal Relevance for a speaker to use an ambiguous utterance and expect the ambiguity to be recognized (much less interpreted). Yet there are classes of implicatures which depend on the addressee’s recognizing two (or more) distinct propositions which the speaker could rationally have intended to communicate in that situation with that utterance. On S&W’s account, however, the addressee can never get past the first one, so it is futile for Mozart, in *Amadeus*, to intend (indifferently) to either compliment, insult, or frustrate Salieri by responding to Salieri’s expectation of approbation by saying (1):

1. I never thought music like that was possible.

Except for the scope of the negative, irrelevant here, the sentence is syntactically and semantically unambiguous. Nonetheless, Salieri can never know whether Mozart intended to convey that the music struck him as wonderful beyond all imagination (likely in the hearer’s context of expecting approbation) or as terrible beyond all imagination (likely for Mozart, as

portrayed in the film). S&W say that “if two essentially different interpretations seem to come simultaneously to the mind of the addressee, and they are both consistent with the principle of relevance . . . the addressee will be unable to decide what the informative intention was, and communication will fail” (p. 169). But if Salieri cannot tell whether Mozart meant his remark as praise or insult, the communication succeeds rather than fails, because Salieri does not conclude that Mozart has failed to communicate: Either he believes that Mozart meant one or the other (and he’ll be hard put to find out which), or he infers that Mozart *meant* for him not to be able to tell. Either way, Mozart wins, *by communicating* ambiguously.

It also follows from S&W’s account that silence cannot be an ostensive stimulus for ostensive–inferential communication. The definition of ostensive stimulus does not rule out silence as an ostensive stimulus, however, because all it requires is that the stimulus attract the audience’s attention and focus it on the communicator’s intentions (p. 153). Remaining silent when one is expected to speak can be quite eloquent. Thus when David (in the television program “Moonlighting”) says (2) to Maddy, and Maddy “responds” by saying nothing, David is so affected by her communicating *thereby* that she does not believe in God, that his response is to entreat her *not to speak*:

2. David: What about God? God defies explanation, and you believe in Him.
3. Maddy:
4. David: Oh, no! Don’t say it!

Similarly, when A says to B, “You believe me, don’t you?” and B, looking directly at A, says nothing, A knows from the fact that B has not responded in the affirmative that the answer is “No.”

Such silences are meaningful in the context of expected behavior which they fail to satisfy. But how could remaining silent be the *most relevant* ostensive stimulus Maddy or B could have used? Surely B would know that it would be easier for A to interpret “No” than to determine his attitude from his not saying “Yes,” and surely Maddy would know that it would be easier for David to interpret the assertion “I don’t believe in God” directly than to infer it from her *not* saying, “All right, I guess I can believe in something that defies explanation.” Are there contextual effects of implicating here which are so large that they outweigh the presumably great effort required to make the inference from silence? Silence here has contextual effects, on the order of “I don’t want to make you feel bad by saying this in so many words,” which the direct answer lacks; but do they outweigh the processing effort required to harvest them? How could you tell? Does an aardvark weigh more than a sonnet rhymes?

#### NOTES

1. S&W’s interpretation of Grice’s views (pp. 161–62 and elsewhere) strike us as unrepresentative. Whether their characterization of other scholars’ views is more accurate is impossible to tell because the views are often attributed to unidentified “pragmatists.”

2. An interpretation is considered “consistent with the principle of relevance if and only if a rational communicator might have expected it to be optimally relevant to the addressee” (p. 167).

## Inference and information

Philip Pettit

Research School of Social Sciences, Australian National University, Canberra, Australia 2601

If we assume that the meanings of words can be treated as something given, then it is useful to distinguish three levels in the linguistic determination of further meaning:

1. The meaning of the words in combination determines the



meaning of the type-sentence: the meaning, for example, of "That is the largest cat," taken in abstraction from context.

2. The meaning of the type-sentence in context determines the meaning of the token-sentence: the meaning of an occurrence of "That is the largest cat" when it is contextually clear that the speaker is talking of caterpillar trucks, that he is, in fact, referring to one in a shop window, and that the reference group within which it is largest is that of the caterpillars on sale in the shop.

3. The meaning of the token-sentence on the lips of the communicator determines the meaning of the utterance: The meaning in the case of our example might be that the caterpillar identified is the one the speaker needs to buy.

The interpretation of meaning requires insight at each of these three levels. The interpreter must be in a position to tell the character of the type-sentence, the content of the token-sentence, and the message of the utterance. The example given has been chosen so that these come apart, but in other cases they may not do so. We should be open to the possibility that the meaning of a nonambiguous, nonindexical type-sentence may yield the meaning of the token-sentence directly, as well as that the meaning of the token-sentence may be nothing less than what the speaker intends to communicate.

The demarcation of these levels of interpretation enables me to voice a reservation about Sperber & Wilson's (S&W's) otherwise admirable *Relevance*. The reservation is that they overemphasise the distinction between level-one interpretation and interpretation at the other levels but underemphasise the difference between interpretation at level two and level three.

First the overemphasis. S&W see the difference between level-one interpretation and interpretation at the other levels as that between a decoding and an inferential process. As they put it, "An *inferential process* starts from a set of premises and results in a set of conclusions which follow logically from, or are at least warranted by, the premises. A *decoding process* starts from a signal and results in the recovery of a message which is associated to the signal by an underlying code" (pp. 12–13).

This statement of the difference is not satisfactory, because a decoding process, as described, can always be cast as inferential. It starts from premises to the effect that a signal of a certain sort has been emitted and that such and such rules of codification are relevant, and it results in the conclusion that a particular message is intended. The sort of conclusion offered at the first of our three levels of linguistic interpretation is, in an example offered by the authors themselves: "Mary has uttered the sentence 'It will get cold'" (p. 177).

In relation to this consideration, it may be said that all that is registered psychologically is a mode of perception: The audience hears the sounds as an instance of the sentence-type "It will get cold"; it does not record the conclusion "Mary has uttered the sentence 'It will get cold.'" But that proves nothing. It is equally true that the second level of interpretation may generate only a mode of perceiving the remark and not a conclusion of the appropriate sort. It may lead the audience to hear the sentence-type as a way of making a certain remark without inducing any explicit conclusion of the form "Mary says that it will get cold."

The notion of inference is a tricky one that, I believe, causes unnecessary problems for S&W. As it is used in cognitive psychology it picks out a process that need not be conscious, personally controlled, or explicitly represented. S&W are attentive to the fact that inference need not be conscious or personally controlled, admitting that it may have an "automatic, unconscious, reflex quality" (p. 83). They do not advert to the fact that it may not be explicitly represented: that is, that a process may be interpretable as inferential even though it does not explicitly represent inferential steps (Cummins 1983). [See also Stabler: "How Are Grammars Represented?" *BBS* 6(3) 1983.] But S&W would surely find that fact congenial, because they do not think that the probability attached to an assumption

is always explicitly represented (pp. 78–79). And yet they fail to recognise that, once these qualifications are entered, then decoding would seem to count as a clear instance of inference.

The difference between linguistic interpretation at the first level in our schema and interpretation at the other levels is a difference between types of inference. At the first level the inference, we might say, is code-based; at the other levels it is based on facts about the context or communicator. It may also be, as S&W allege, that the first-level inference is conducted by a peripheral processing system, the other sorts of inference by a central one (p. 71). The point does not seem important at this abstract level of analysis.

So much for the overemphasis with which I charge the authors. The underemphasis that I discern appears in their discussion of the second and third levels of linguistic interpretation. They suggest that although differences are involved, the divide between these levels is not of any great import. That, I believe, is a serious mistake.

The suggestion comes early on when S&W set up a contrast between the case in which someone provides direct evidence of an assumption that he wishes to make manifest to his audience and the case in which he manifests the assumption by other means. Peter leans back on the park bench, allowing Mary to see directly that the unwelcome William is approaching and only incidentally, if at all, to recognise that he intends her to see William (p. 48). The case is designed to contrast with one in which he says that William is coming rather than showing it in this way.

The important question is how the contrast is supposed to go. S&W seize on the fact that in the showing case direct evidence of the assumption is provided; they suggest that in the saying case the evidence must be indirect. "When a coded signal, or any other arbitrary piece of behaviour, is used ostensively, the evidence displayed bears directly on the individual's intention, and only indirectly on the basic layer of information that she intends to make manifest" (p. 53). The idea, familiar from the work of Grice, is that by uttering a sentence like "William is coming" the speaker manifests his intention to manifest the assumed fact that William is coming and thereby hopes to manifest that fact itself.

But an alternative picture is available. This is that, just as what Mary sees directly presents the assumption which Peter wishes to make manifest, so the sentence-token "William is coming" – the sentence-type in context – directly presents or, better, represents that assumption. It may not do so on the basis of decoding alone; it may require both context-based and code-based inference. But the important possibility to register is that it conveys that assumption immediately and not by first presenting evidence of the speaker's intention to manifest the assumption. As in the showing case, the communicator may indeed present evidence of that intention; but that he does so is incidental to the success of his basic communicative enterprise.

The alternative picture clearly needs to be developed to account for a number of facts which I ignore here, as for example that sentence-tokens may be nonassertoric (McDowell 1980). There are a number of reasons why it is surprising that S&W do not consider this picture. One is that they refer to McDowell's book, which explicitly introduces it. Another is that, recognising that bees communicate linguistically other than by identifying one another's intentions, they ought to be open to the possibility that humans can do so too (pp. 5, 174). And a third reason for surprise is that, although bees may not have to exercise context-based inference, that difference should not block the possibility mentioned since, by S&W's own account, the relevant features of context invoked in such inferences do not concern the speaker's intentions (pp. 183–93).

The alternative picture is phenomenologically and methodologically attractive – phenomenologically because we do seem to identify the meaning of sentence-tokens without any reference, conscious or unconscious, to the intentions of speak-



ers; methodologically because the picture puts us on a continuum with nonhuman animals. We can see the ability to recognise sentence-token meaning as emerging prior to any capacity for divining one another's intentions. [See also Dennett: "Intentional Systems in Cognitive Ethology" *BBS* 6(3) 1983.]

The picture, finally, has a number of more incidental attractions. It suggests a short line with the overtness problem: Communication at this basic level will be overt just so long as the speaker lacks any intentions to be covert; it does not require the presence of positive intentions of transparency (Evans & McDowell 1976). The picture also enables us to define a distinctively linguistic form of belief as the disposition to assent to certain sentence-tokens; this supplements the more basic notion of belief simply as a disposition to act as if certain things were so. Lastly, the picture opens up the possibility that, if context is understood broadly enough, then metaphorical content can be seen as determined at level two rather than level three, so that there may be no literal way of thinking certain metaphorically expressed thoughts (Goodman 1968).

Nothing would be lost to the basic thesis of *Relevance* if the authors were to endorse this alternative picture; it would remain the case that considerations of relevance determine meaning of the sort that appears at level two. The only difference would be that the relevant relevance would usually be the sort involved in the cognitive processing of contextual phenomena, not the kind associated distinctively with the divination of speaker's intentions (pp. 151–55).

## The relevance of *Relevance* for fiction

Anne Reboul

Department of French, University of Geneva, 1205 Geneva, Switzerland

I would like to outline briefly the way in which the framework devised by Sperber & Wilson (S&W) can be helpful to students of fiction in general, and especially to students of literary fiction.

There have been two main approaches to fiction. On the logical side, as a response to the fact that fictional utterances are false, the adoption of a possible-worlds semantics has been proposed. On the pragmatic side, as a response to the loss of normal illocutionary force in fictional utterances, the idea has been advanced that there is a specific act (whose illocutionary status is rather obscure) which would enable the speaker to produce an utterance without intending to bring about any of its usual illocutionary consequences. However, it seems that neither of these theories is entirely satisfactory. While possible-worlds semantics offers a solution to certain types of fictional discourse, it offers no solution to the kind of inconsistent or paradoxical utterances which can appear in fictional discourse. The pragmatic approach has been on the whole vague (little has been said about how the act performed in fiction works), and at best metaphorical.

Relevance theory offers a rather different approach. S&W are interested not only in the sense of utterances, but in the overall interpretation process by which contextual implications of utterances can be recovered. There seems to be no special reason for thinking that fictional utterances are interpreted in a fundamentally different way from utterances taken to be true. Whether true or false, any grammatical utterance encodes a logical form enabling the hearer to construct a context against which the utterance can be interpreted, as S&W's analysis of metaphor has shown. And as they also show, in any utterance, fictional or otherwise, the principle of relevance, with its twin factors of contextual effects and processing effort, can shed interesting new light on style.

This does not mean that no distinction should be drawn between fictional and ordinary utterances. One specific prob-

lem raised by fictional utterances is that the objects to which they purport to refer do not exist. This does not, of course, prevent fictional names from being associated with mental concepts in S&W's sense. It has been suggested in Reboul (1986) that the concepts corresponding to definite descriptions or proper names in fiction should contain in their encyclopedic entries, apart from the information provided by the fictional discourse itself, an additional item specifying their origin (i.e., for Hamlet, the play *Hamlet*), whereas the entry corresponding to that origin (i.e., the concept attached to the drama *Hamlet*) should contain information specifying that it has no extension.

Another question which any theory of fiction should answer is why people are prepared to spend time reading or listening to fiction, even though they know that it is false. Neither possible-worlds semantics nor the traditional pragmatic approach considers this question. There is, however, an answer suggested by S&W's framework and argued by Reboul (1986): it is that fictional discourse should be looked on as very near to metaphor. On that view, fictional discourse would be an interpretation of a thought (or thoughts) of the speaker. Like metaphor, it need not purport to represent an actual state of affairs as long as it achieves optimal relevance: that is, as long as it gives rise to sufficient effects which describe actual states of affairs (the "message" of the work) for the minimum necessary processing effort.

It could be argued against this view that one can find in fictional discourse all kinds of utterances which, in "ordinary" discourse, would be considered descriptive or interpretive in the usual way. However, within S&W's framework, hearers or readers need not rely only on the form of an utterance to determine whether that utterance is an interpretation in the same sense as metaphor, or whether it is an interpretation or a description in the usual sense: As always they consider which of these possibilities would yield an interpretation consistent with the principle of relevance.

It appears, at least to me, that the approach the most likely to enhance the relevance of fictional utterances might be to consider them as interpretations of thoughts of the speaker in much the same way as metaphor, and that this is the most productive way of approaching fiction.

## Literalness and other pragmatic principles

François Recanati

Centre National de la Recherche Scientifique, 15 quai Anatole France, 75007 Paris, France

Sperber & Wilson's (S&W's) achievements in *Relevance* are most impressive. First, they offer a new picture of human cognition, according to which the mind processes centrally available information in such a way as to obtain the greatest improvement of its representation of the world at the smallest processing cost. Both the information the mind chooses to process and the manner in which it is processed (with more or less extensive appeal to background knowledge) are said to be determined by this search for "relevance." Second, they show that this view of human cognition provides the foundation for a theory of human communication, both linguistic and non-linguistic. Third, they suit the action to the word and give a detailed theory of linguistic communication. This theory I find very appealing. It is more explicit than most theories on the market; it accounts for hitherto neglected phenomena such as the communication of feelings and vague thoughts; above all, it is the only theory that both properly acknowledges the linguistic underdetermination of what is said and yet accounts for the hearer's ability to select a unique interpretation. Nevertheless, if we abstract from the unified, cognitively based framework they provide for pragmatics, it seems to me that S&W's theory is

not as different as they claim from those of other pragmaticians in the Gricean tradition.

Consider, first, the alleged difference between the principle of relevance and Grice's cooperative principle. According to S&W:

There is . . . a radical difference between the principle of relevance and Grice's maxims. Grice's cooperative principle and maxims are norms. . . . Communicators generally keep to the norms, but may also violate them. . . . The principle of relevance, by contrast, is a generalisation. . . . The principle of relevance applies without exception: Every act of ostensive communication communicates a presumption of relevance. (Précis, sect. 3.3, para. 7–8)

I do not think there is so radical a difference. In the Gricean framework as I interpret it, a communicator is expected not only to conform to the cooperative principle but, by the very act of communicating, presents himself as conforming to the cooperative principle. (If this were not so, the conversational implicatures – that is, the hypotheses that are needed to reconcile the utterance with the presumption that the cooperative principle is respected – would not be communicated, as they are, but simply induced in the hearer.) This is not very far from the notion that every act of communication communicates that the cooperative principle is being respected. As for the cooperative principle itself, formulating it exclusively in terms of relevance would not seriously affect the spirit of Grice's theory.

Another example of what is perhaps S&W's tendency to overstress the difference between their view and those of other pragmaticians is their statement to the effect that "most pragmaticians . . . see the inferential tier of comprehension as governed by a variety of specialized rules constituting a kind of pragmatic 'module'" (Précis, sect. 4, para. 2). This statement may be true of (some) formal pragmaticians, but certainly is not true of inferential pragmaticians in the Gricean tradition. Like S&W, these theoreticians take pragmatics to be concerned with the interaction of linguistic competence with world-knowledge and general intelligence. It is true that pragmaticians generally do not refrain from positing pragmatic rules and principles; but they do not believe that doing so is inconsistent with the view of pragmatics as nonmodular. In general, the pragmatic principles put forward in the literature are only more or less specific applications of higher-level principles such as Grice's cooperative principle, to which they ultimately reduce. Do S&W think that a pragmatic theory that incorporates a version of Grice's maxims is thereby committed to a modular view of pragmatics? That would show misunderstanding of the function and status of the maxims in a pragmatic theory.

Independent of Grice's maxims, it has often been assumed that there is a "principle of literalness" at work in communication. With respect to such a principle, S&W's criticism may well be justified. However, in this case as in others, what is relevant to pragmaticians is not so much the principle itself as the presumption that the principle is respected – that is, the presumption that a speaker is speaking literally. Now, to say that such a presumption plays a role in the comprehension process is simply to say that the literal interpretation of an utterance is considered first. This, of course, is consistent with the non-modular view. What is to be criticized, therefore, should not be the substance of the theory but the way it is formulated.

As far as the principle of literalness is concerned, S&W reject not only the formulation but also the substance of the claim: They do not believe that the literal interpretation comes first. If they are right, this is an important result. However, S&W's view is difficult to assess because, in discussing literalness, they have recourse to an extra piece of theoretical machinery, expounded in section 4.5 (Précis). I do not understand why they do not keep to the simple (unextended) relevance-theoretical framework; accordingly, I suggest the following restatement of their view concerning irony, metaphor, and speech acts:

(A) In verbal communication, S's utterance of a declarative sentence manifests his intention to make manifest something that bears a certain relation  $R^*$  to a proposition coming under the semantic

potential of the sentence. This something – the *communicatum* – may be a proposition expressible by the sentence (then  $R^*$  is identity, and the speech act performed is an ordinary assertion), or an expansion of such a proposition (then  $R^*$  is the relation of embedding; this covers both irony and speech acts other than ordinary assertions), or a proper subset of the entailments of such a proposition (this covers metaphor and loose talk). The hearer's inferential task is to identify the relevant *communicatum*  $R^*(P^*)$ . Starting with the meaning of the sentence, he searches for a proposition  $P^*$  such that (i)  $P^*$  is expressible by the sentence and (ii) there is a relation  $R^*$  such that, if S intends to make  $R^*(P^*)$  manifest, then his utterance is consistent with the presumption of relevance.

The claim that there is a presumption of literalness is not inconsistent with S&W's view thus stated. According to (A), to test a candidate  $P$  for the status of "proposition expressed" is to search for an  $R^*(P)$  such that the hypothesis that  $R^*(P)$  is communicated is consistent with the principle of relevance. In this framework, the claim that there is a presumption of literalness says only this: The first candidate for the status of  $R^*(P)$  to be considered, the most accessible one given that one is testing a definite proposition  $P$ , is the proposition  $P$  itself (i.e.,  $R^*(P)$  with  $R^* = \text{identity}$ ). It is only when the hypothesis that  $P$  is being communicated is not consistent with the principle of relevance that  $H$  goes on and looks for a relation  $R^*$  different from identity.

## Rationality as an explanation of language?

Stuart J. Russell

Computer Science Division, University of California, Berkeley, Calif. 94720

Sperber & Wilson (S&W) are to be congratulated for their introduction to the problem of communication, and especially for the eminent common sense of their position that the ostensive-inferential model provides a more acceptable general theory than does the code model. Their goal of explicating the cognitive effects of an utterance on the hearer, as opposed to isolating its "meaning," does away with a lot of pointless arguments. As a researcher in artificial intelligence with a somewhat formal bent, I began reading with some trepidation but was thoroughly reassured by their criticism of existing pragmatic theories for choosing interpretations (p. 37): "What they fail to show is that . . . an equally convincing justification could not have been given for some other interpretation that was not in fact chosen."

In other words, S&W are looking for a generative theory of pragmatics. Analysis of cognitive processes as rational behaviour, given sets of beliefs and desires, is often a reasonable first-order approximation; the task is then to establish exactly what are the goals of communicating agents and what knowledge they have available or can assume in their processing, and finally to show that these premises can reasonably be attributed to the agents and are sufficient to generate the interpretive behaviours that are actually found in humans. Unfortunately, the machinery proposed in the theory of relevance falls well short of fulfilling these goals, in several ways.

S&W's first contribution to a new theory is the idea of mutual manifestness, intended to replace mutual knowledge as a means of making inferences about the goals and beliefs of others. Mutual knowledge, which is correctly pinpointed as a *sine qua non* of error-free coded communication, is rejected on the basis of being an unattainable idealization, requiring an infinity of inferences. Pure coded communication is therefore rejected also. This rejection, however, seems to be based on two false premises: the assumption that knowledge has to be certain and the assumption that mutual knowledge requires an infinite chain of explicitly held assumptions. The formulation of Halpern and Moses (1985) gives a definition of mutual knowledge  $M_p$  of proposition  $p$  as

$$Mp \equiv \bigwedge K_i p \wedge K_i Mp.$$

According to this definition, everyone knows  $p$  and everyone knows it is mutually known. Under normal communicative circumstances it is reasonable to commit to a belief in mutual knowledge of  $p$ , which can be done in finite time. Three things make one suspicious of S&W's rejection of mutual knowledge: First, coded communication does, of course, occur; second, it is later admitted that communication need not, indeed cannot, be guaranteed to succeed, even though this requirement was imposed on coding; and third, since manifestness is only a dispositional property, in S&W's definition, it is hard to see how mutual manifestness allows the essential inferences to be made about the actual cognitive states of others. The concept seems to suffer the drawbacks of Hintikka's (1962) deductive-closure model of knowledge without the benefits of an explicit model such as that of Konolige (1986).

Fortunately, the detour via mutual manifestness can be avoided without undue strain on the rest of the theory of relevance. Perhaps a better reason to reject the code model is simply that the distortions caused by reformulating the process of contextual inference as a sign-message correspondence system render the model practically useless as a generator of insights into complex communication phenomena.

S&W propose two main constructs to flesh out their rational approach: a model of human inference and a definition of relevance, which plays the role of the human cognitive goals in the rational communication process. These cognitive goals are assumed, in some "trivial, speculative remarks," to be the enlargement of the set of explicitly available facts known and the minimization of processing effort. It might be more consistent with a rationalist approach to include the hearer's goals as a constraint on which new facts are inferred, as in much current work in artificial intelligence.

A model of an inference mechanism is required to give substance to claims about which facts become known as a result of communication, and with how much effort. To a reader familiar with artificial intelligence and logic, the proposals for this mechanism put forward in *Relevance* seem somewhat strange. Here are the problems associated with several claims:

*Claim: Inference consists of unfettered forward chaining to termination.* This is absurdly impractical, as actual experience demonstrates (Schank 1979). Control methods must be used to ensure that only useful inference paths are explored.

*Claim: All inference rules are explicitly represented.* Such a scheme cannot succeed, because whatever combines the inference rules with their premises must also be an inference rule, but procedurally represented; moreover, almost all actual inference machines incorporate their inference rules by means of procedures because this is far more efficient, and inference rules never change.

*Claim: All inference is eliminative.* This would disallow the use of commutativity laws and axioms defining inductive domains, for example.

*Claim: Assumption schemas are not complete propositions, yet can be used in inference.* A formal object without a truth value cannot, for example, contradict anything. As admitted by S&W (chap. 2, n. 10) these schemas are really just existential propositions.

*Claim: Assumption strengths, or degrees of confirmation, are not explicitly computed, but are merely produced as the result of nonlogical processing of assumptions; hence confirmation theory is irrelevant.* This seems to be a confusion between the specification of a logical confirmation process and its implementation. By the same token, one could say that a theorem prover using inference procedures rather than explicit inference rules is not a logical system. S&W's rejection of a computational approach to confirmation is especially odd in light of the fairly standard (although shaky in places) probabilistic inference procedure they advocate.

Fortunately, again, these deficiencies can be repaired without undue damage to the rest of *Relevance*. However, it seems odd that little or no reference is made to the considerable work in artificial intelligence on models of plausible inference, both for problem-solving in general and for language understanding in particular. (Kanal and Lemmer [1986] and Norvig [1987] contain worthwhile discussions.) The procedures S&W propose for memory access and context construction are woefully underspecified.

The two principles (it seems there really are two) of relevance and least effort are put to work to explain a large number of comprehension phenomena. It is here that the work cries out for a formal implementation, because in any number of cases we are expected to sit idly by while yet another *ad hoc* rule for estimating effort or for defining and accessing contexts is introduced in order to rationalize a desired interpretation (see, for example, pp. 187, 200, 219–20, 221). Effort, in particular, is far too flexibly defined to be a theoretically adequate tool for a real pragmatic system, and there seems to be no way to obtain more than introspective evidence. Although I will leave it to the linguists to assault with deadly example-sentences, it seems unlikely that the minuscule amounts of effort required by S&W's theory, their motivating picture of mental exhaustion notwithstanding, would play a significant role except in highly time-pressured situations. To place all the burden of selecting correct interpretations on an effort principle, implemented as an ordering on context accessibility, abdicates the responsibility for a theory, passing it on to those studying memory organization, who are already overloaded with ordering constraints passed on by other theories such as the possible-worlds theory of counterfactuals.

In general, one gets an uneasy feeling about the value of some of the explanations proposed in *Relevance*, akin to the feeling experienced when reading teleological explanations in evolution or functionalist explanations in sociology. Although, in principle, one might be able to trace all aspects of language back to their roots in cooperative advantage, the effort may be neither practical nor useful. For example, Quine's (1960) model of how words get their meaning is very interesting philosophically, but it should not be part of an explanation of how language is processed. S&W are apparently happy to accept the existence of convention in the area of grammar and the lexicon, although they reject all other uses of convention as *ad hoc*. Conventions of all kinds are, however, self-justifying, since by their existence they enable more effective cooperation. Even if some of the pragmatic processes involved in communication have a rational explanation, it would seem that our first task is to establish exactly what the processes are, along with what conventions they may rely on. The first job of an Englishman after a transatlantic migration is not to find out the rationale behind a society's choosing one side of the road to drive on rather than the other, but to find out which side of the road to drive on.

## How relevant?

Pieter A. M. Seuren

*Philosophy Institute, Nijmegen University, 6500 HK Nijmegen, Holland*

It is gratifying to see Sperber & Wilson (S&W), in their *Précis of Relevance*, state openly, at the very end, that they are "well aware that the view developed in *Relevance* and summarized here is very speculative and, as it stands, too general directly to determine specific experimental tests or computer simulations." Nowhere earlier had this been said so explicitly. One is now fully justified in concentrating on the internal logical coherence of the view in question, for if a speculative view has any force, it is mainly on account of its internal structure and logical clarity. S&W's critics, including myself, have always concentrated on precisely these issues, and, as it now appears,



rightly so. In my (very negative) review of *Relevance* (Seuren, 1988) I raise a number of central issues having to do with the coherence and consistency of S&W's speculations, to which Wilson and Sperber (1988b) wrote a (rather piqued) reply. I shall now take up some of the objections left unanswered, or inadequately answered, in that reply.

The central theme in S&W's view of relevance is the fact that "there is a gap between the semantic representations of sentences and the thoughts actually communicated by utterances. This gap is filled not by more coding, but by inference" (Précis, sect. 1.1, para. 3). Thoughts are "conceptual representations" (*Relevance*, p. 2). Among what are considered thoughts, *assumptions* occupy an important place: They are "thoughts treated by the individual as representations of the actual world (as opposed to fictions, desires, or representations of representations)" (p. 2).

The relevance of an assumption in a context is said to depend on the extent of the *contextual effects* in that context, as well as on the effort required to process it in that context: More contextual effects and lesser processing effort lead to greater relevance-in-that-context of the assumption in question. One has to take it, therefore, that an utterance *u*, together with an unknown quantity *i* of inferencing in a context *C* yields a set of thoughts (assumptions) *I* actually communicated by *u*. *I* is graded for relevance according to the wealth of contextual effects and the minimality of the processing effort. This grading is not absolute but relative to other possible assumptions communicated by *u* with other inferencings and/or with other contexts. Each pairing of some *u* with some *i* in some *C* is called a *hypothesis* (p. 165), let us say *H*. Each *H* thus determines some set of assumptions *I*, with an associated relevance value *R*. There is then a *Presumption of optimal relevance*, which makes the hearer select the *I* with the highest *R*, at least in some cases (p. 165): "In some cases, [the task of constructing and selecting a hypothesis] is best carried out by listing all the possible hypotheses, comparing them, and choosing the best one." In other cases, however, "it is better carried out by searching for an initial hypothesis, testing it to see if it meets some criterion, accepting it and stopping there if it does, and otherwise repeating the process by searching for a second hypothesis, and so on." One might now wonder what can be meant by "processing effort," in that the processing of an utterance involves the adequate selection of a context (pp. 137–38): What if *u* yields a very high *R* (lots of contextual effects, at very little processing cost) at hypothesis 358? Would not the processing cost needed to get to H358 outweigh the saving at H358? Wilks (1986, p. 273) observes that this seems to make nonsense of the condition of minimal processing effort. S&W's answer (p. 167) is that, in fact, the interpretation to be selected is "the first one tested and found to be consistent with the principle [of relevance]. . . . The important point is that, given the cognitive environment, given the initial context, and given the stimulus, some hypotheses are more accessible than others, and this means that they require less processing effort." Had the communicator intended the hearer to settle for a less accessible hypothesis, he should have used an utterance that saves the addressee the effort of first accessing more accessible hypotheses (p. 168).

All this raises a number of questions. First, S&W's account requires that, given a context (including a "cognitive environment"), a given utterance *u* determines the order of accessibility of possible interpretative hypotheses. Apart from the empirical problem raised by this requirement (we have promised to let that aspect rest), there is the fact that the function involved is left totally unspecified, crucial though it is in the total framework. Having said (p. 167) that *u*, in virtue of its grammatical and lexical form, gives immediate access to a highly determinate set of concepts, assembling them into a logical form directly usable as an assumption schema, S&W content themselves with stating: "The context provides ways of completing these assumption schemas into full hypotheses." This conclusion is presented as a

fact, and not as a serious white area on a map whose filling in is doubtful. Unfortunately, this type of procedure is typical for *Relevance* (and its Précis): Very often proper analysis and argument have to give way to rhetorical suggestions and demonstration by way of examples. Yet the result is put forward as solid truth. The book would have gained a lot if it had been written less apodictically.

It would also have gained a lot if it had been written more clearly. What one finds is an almost inextricable maze of definitions couched in often unnecessarily opaque and abstruse terminology. One fails to find clear statements of the computational aspects (functions) involved or invoked, or of the rules or principles needed to make the less determinate parts of the system work. This is perhaps best illustrated by holding up for close inspection S&W's notion of contextual effect (CE). We remember that without CEs there is no relevance. What, then, is a CE? It is meant to be an alteration of a given context resulting from the addition of a new utterance (not counting the purely incremental effect of the utterance). The context *C* can be modified in three possible ways (p. 117): Some assumption may be removed from *C*, its "strength" may be modified, or *C*, enriched with *u*, may now possess a richer set of entailments than either *C* or *u* alone. In the last case we have a so-called *contextual implication* (pp. 107–8). Moreover, the strength of any assumption depends on the evidence available for its truth (p. 109); and the highest degree of strength, one infers from p. 113, is certainty. Strength adheres not only to the contents of the speaker's message but also to the listener's recognition of the speaker's intention as part of the set of assumptions transmitted. When a conclusion is drawn, its strength will depend on the relative strengths of the input assumptions (pp. 109–16). But S&W's text is rather verbose but sketchy and imprecise on this point: No precise procedures or formalisms are provided. Moreover, if strengthening helps comprehension (by raising *R*), how can weakening do the same? According to S&W, if *u* is contradictory with some element *e* in *C*, then the weaker of the two will be erased from *C* (p. 115). If *e* is the weaker one and is eliminated, we have a CE, but if *u*, the new utterance, has to be kept out, there is no CE. Erasure of an *e* can be considered the most extreme case of weakening, and this, too, helps comprehension! Would not some further clarification be in order here?

All in all, this centrally important notion in S&W's framework remains without a precise definition. In Wilson & Sperber (1988b) it is claimed that *Relevance* (p. 260, n. 26) does offer a formal definition. In this footnote a "more formal characterisation" is promised. What follows is my rephrasing of the relevant text in more easily accessible terminology and symbolism. Let *C* stand for context, *P* for newly added propositions ("assumptions");  $E_C$  and  $E_P$  represent the set of conclusions derivable from *C* and *P*, respectively. The contextualisation of *P* in *C* has no CE just in case (i)  $E_C \subset E_{C \cup P}$ , and (ii)  $E_{C \cup P}$  minus  $E_C \subset E_P$ . If these two conditions are not both met, there is some CE. But notice that condition (i) is trivially met in all possible cases, since whatever follows from a set of propositions *C* also follows from *C* enriched with any set of propositions. So the only condition is (ii): For there to be a CE there must be a conclusion in  $E_{C \cup P}$  which is neither in  $E_C$  nor in  $E_P$ . But this is precisely the definition given for *contextual implication* (pp. 107–8), and thus, what is defined in note 26 is an unnecessarily complex rephrasing of what had already been given as a definition for *contextual implication*, and does not define *contextual effect*. The elements of "strength" and "erasure through contradiction" are still undefined. It must be added that S&W also say in note 26: "Let two assumptions with the same content but with different strengths count as two different assumptions." But is one to understand that if only strengthening occurs (whatever that may be), then, due to a mysterious change in identity of the strengthened assumption set, there are suddenly also contextual implications which, without the strengthening, would

not have been there? If so, the reader might expect considerably more explanation.

This leads us to S&W's notion of *contextual implication*: A set of assumptions *P* contextually implies an assumption *q* in the context *C* if and only if (i) the union of *P* and *C* nontrivially implies *q* and (ii) neither *P* nor *C* nontrivially implies *q*. This, as we have seen, is identical with the definition of contextual effect (p. 260), except for the specification that the deduction must be "non-trivial." (I take it that "non-trivial implication" is also what is meant in that definition. Otherwise I cannot understand what is meant.) What, then, is "non-trivial implication"? For this we must go to p. 97: A set of assumptions *P* logically and nontrivially implies an assumption *q* if and only if, when *P* is the set of initial theses in a derivation involving only elimination rules, *q* belongs to the set of final theses. By "derivation" S&W mean a formal, logically deductive procedure whereby logical consequences (entailments) are computed. S&W maintain, no doubt correctly, that standard first order logic is both too poor and too rich for the proper analysis of natural language entailments: It is too poor because it fails to capture many if not most intuitive entailments, and it is too rich because some formal entailments in logic are counterintuitive. In order to remedy the poverty, S&W invoke a "concept logic," the details of which are not provided. Against the excessive logical power they posit that in language and cognition only elimination rules are used and never introduction rules. This in itself has extremely awkward consequences, since it rules out the conclusion that two people came in if it is indicated that Henry and Jack came in – to mention just one example. Moreover, the notion of elimination rule is left opaque (for details see Seuren, 1988). The nonstandard logic invoked by S&W thus seems not to exist in the sense of being formally available. S&W further claim that, in their logic, the set of entailments associated with any *P* is finite. No evidence is offered for this bold thesis, yet S&W state that "The process [of derivation or deduction] applies to all initial and derived theses until no further deductions are possible" (p. 95).

Let us now turn back to the definition of nontrivial implication as paraphrased above (from pp. 107–8). What is meant by "the set of final theses"? Note that this term is simply used in this definition but nowhere explained or defined. In Seuren (1988) I took it that, in accordance with current terminology in the theory of algorithms, it should refer to the set of ultimate conclusions from which no further deductions are possible. Wilson & Sperber (1988b), however, indicate that this is not the intended interpretation: The term must be taken to refer to the union of the initial and the derived theses (assumptions). Not only does this terminology seem sloppy, but "the set of final theses" is still not defined, and hence neither are the notions of nontrivial implication, contextual implication, and contextual effect.

This is only one of many similar tangles in the book. What, then, remains of S&W's speculative view? Very little, in my opinion. I know of no other attempt of the same broad scope and proportions to explain the complexities of linguistic and non-linguistic comprehension of messages. Perusing both *Relevance* and its *BBS Précis* calls to mind why no alternative approaches have been attempted: The enterprise seems altogether too ambitious, for the time being, to be realistically undertaken. In my view it is much more realistic and fruitful to let the wide and general issue rest for a while, in order to concentrate on detailed empirical analyses of more restricted problem areas. The empirical knowledge thus gathered will prove highly relevant in constraining and shaping speculative hypotheses about the broad kinds of issues that S&W have the courage to tackle. There is plenty of room within the bounds of the empirically accessible for fruitful theory to develop.

## The information needed for inference

Carlota S. Smith

Department of Linguistics, University of Texas at Austin, Austin, Tex. 78712

Sperber & Wilson (S&W) argue that verbal communication involves both code and inference. In making inferences people use general abilities and knowledge rather than the specifics of a code; and the fact that people revise and abandon interpretations is strong evidence in favor of S&W's position. Writers often exploit this flexibility, for example, by using unreliable narrators in fiction. Readers must gradually revise their interpretation of a text as they lose faith in the narrator.

The many inferences in S&W's account of communication require several types of information, principles, representations, and computations. The principle of relevance seeks an optimal balance between effort and effect. As an essential first step toward understanding how the principle might actually operate, I wish to consider the information required for inference at the various stages of understanding. The variety makes it somewhat difficult to see how a single principle might apply.

There are three main stages in understanding, according to the programme of S&W: decoding an utterance, arriving at its propositional form, and assessing its import. The last two stages require inference. Consider utterance (b) below, uttered by B in the context of A's uttering (a). A's task is to understand (b).

A: (a) Susan has thalassemia.

B: (b) She is getting married to Bill.

Assume that A, having decoded (b), is in a position to infer its propositional form, or explicatures. The explicatures are based in part on the logical form of (b). To work them out in detail we would need a representation of logical form that could be used in inferences concerning ambiguity resolution, reference assignment, completion or enrichment, and determination of propositional attitude.

*Ambiguity*: A must decide between the interpretations decoded for an ambiguous utterance, choosing the one most compatible with the situation. Since (b) is not ambiguous, take S&W's example *The child left a straw in the glass*. To resolve the lexical ambiguity the receiver consults encyclopedic memory, which contains stereotypical scenarios that chunk information and are relatively accessible; such a scenario, once found, would give the "drinking tube" interpretation. Memory also contains the "piece of hay" interpretation, and presumably other information about straw as well. But to work out the effort required to consult encyclopedic memory, we need to know how it is organized. Specifically, in what sense is a chunk accessible? What steps are required? What choices are available? In addition, we need to know how situations are represented and how they affect the search of memory. Note that if the sentence above were uttered during a country picnic the second interpretation might be preferred.

*Reference assignment*: A must determine the reference of the pronoun *she*. Suppose that A finds no third person in the environment and therefore discards the possibility of a deictic interpretation; A then chooses the anaphoric and looks for an antecedent, finding in (a) a noun phrase of the correct gender and number. This information is quite different from that required for ambiguity resolution, as are the calculations involved. To determine the reference of *Bill*, A draws on knowledge of people familiar to both A and B; such knowledge must be represented and computed.

*Completion*: The present progressive in (b) is vague and must be filled out by inference. It may refer to a marriage in progress or a future marriage. Suppose that A, consulting the immediate environment, fails to find a wedding in progress. But perhaps the wedding is going on somewhere else: A must search memory, perhaps expanding the context, for information on this point.



No ongoing wedding is found and A prefers the future interpretation. To work out these inferences we need to know how information about the environment and is represented and stored in memory.

*Propositional attitude:* A must decide whether (b) is an ordinary assertion. A does this by considering the manifest intentions of B in uttering (b) and computing the implications in order of accessibility. I do not know how A approaches this task; let us assume that (b) is inferred to be an assertion.

The explicatures of (b) have now been worked out and A must determine its implicatures. To do this A must choose the context. Although S&W argue that this move is essential to limit the domain of inference, they do not address it directly: In their examples the context is already chosen. A may choose to expand the context by considering the immediate physical environment, earlier utterances, or encyclopedic memory. Note that each involves different types of information and different representations. Suppose that A searches memory for *getting married* and *thalassemia* and arrives at the assumptions (i) People who are getting married should consult a doctor about hereditary risks to children and (ii) Two people with thalassemia should be warned against having children. These will provide the context for determining the implicatures of (b). But how does A arrive at them? Are they retrieved or constructed, for instance? There may be several scenarios in memory for *getting married*. Is it easier to construct (i) than to find it? Questions like this must be answered in addition to the earlier questions raised about memory. (The final stage of inference involves the processing of assumptions; I must ignore this here, for lack of space.)

At various stages of understanding A searches memory, the immediate environment, and the context for information. This information is used to retrieve and construct assumptions that are used in inference. S&W do not present detailed accounts of memory, representation or logical form; they also decline to engage in computation. This makes it quite impossible to work out how information is accessed and how assumptions are constructed; in short, to see how the processes of inference work. The principle of relevance may be a general guiding principle that is realized differently in different areas; or it may be a single principle that drives inference. One looks forward to the presentation of detailed models that will allow these issues to be addressed.

## On interpreting “interpretive use”

N. V. Smith

*Department of Phonetics and Linguistics, University College London, London WC1E 6BT, England*

One major achievement of Sperber & Wilson’s (S&W’s) theory of relevance is that it provides conceptual categories and distinctions that bear fruit in areas distinct from those for which they were devised. Once one has absorbed *Relevance* it is hard not to see its relevance ubiquitously: from language acquisition (cf. Smith 1987) to art appreciation. For lack of space, I want here to focus on one subpart of the theory (the contrast between “descriptive” and “interpretive” use), illustrating its application in a number of areas and suggesting that the diagram offered as Figure 1 in the *Précis* (fig. 3, p. 232) encapsulates their claims about how the distinction should be both modified and extended.

It should first be noted that “Interpretation” is used in two distinct, if related, senses. S&W are, indeed, at pains to make this explicit; but it would have been helpful if the interpretive relation between the propositional form of an utterance and a thought of the speaker (the relation between the top two boxes in fig. 1) could have been labelled something like “(partial) characterization” or, as on p. 192, “incompletely represented.” (The partial nature of this relation is a function of the aniso-

morphism between sentences of natural language and sentences in the language of thought: an important point in its own right and one that S&W make with characteristic clarity and emphasis). This terminological quibble is prompted by the observation that their characterization of tropes comes to the illuminating conclusion that the classical figures do not form a natural class, but are distinct, at least in part: Metaphor, for instance, is an interpretive relation of the kind just mentioned, whereas irony is an interpretive relation of the kind indicated in the lower left-hand corner of the diagram: namely, an interpretation of an attributed thought. More specifically, “interpretive use” refers to the case in which “the thought of the speaker which is interpreted by the utterance is itself an interpretation . . . of a thought of someone other than the speaker” p. 238.

This contrast is important and original, but I have two worries. Prototypical exemplifications of the four categories at the bottom of the diagram are provided (see p. 231) by irony (a), interrogatives (b), assertions (c), and requests (d), with (a) and (b) being interpretive and (c) and (d) descriptive. There is a temptation to infer that these categories are typically mutually exclusive; but while this may be close to true of (b), (c) and (d), it is clear that irony cross-classifies each of the other categories. Thus we can have – as S&W are aware – ironical assertions, ironical requests, ironical questions, and so on. This suggests that the class of ironical (and perhaps “echoic” uses more generally) should be orthogonal to the interpretive/descriptive dichotomy. The category “interpretation of an attributed thought” is better exemplified by the kind of interpretive utterance characterised in some languages by so-called “hearsay” particles. These are particles (as, e.g., *(ha)ti* in Hixkaryana – cf. Derbyshire 1985; p. 127) appended to sentences to indicate that the speaker is disclaiming responsibility for the reliability of the utterance concerned and is attributing this responsibility to some other person (specified or not). In an interesting paper couched within the relevance-theoretic framework, Blass (1987) has argued that the hearsay particle *re* in Sissala “might be more explanatorily described as a marker of interpretive use” in that it characterises sentences involving “free indirect style,” the attribution of thoughts to other people, and, significantly, irony. It is clear that her suggestion would generalise to the Hixkaryana examples and to such cases as the Turkish “evidential” expressed by the particle *-miş* which, according to Slobin and Aksu (1982), “carries modal functions of inference and hearsay, and is pragmatically extended to expressions of surprise, irony and compliments” (p. 185; cf. also p. 195). It is clear that ironical utterances are interpretive; it seems equally clear that they are not interpretive in the same sense as either of the other uses of this term, and should perhaps be characterised not as “second-degree interpretations” (p. 238) but as third-degree interpretations: a stage at which terminological problems border on the substantive.

The diagram summarising descriptive and interpretive dimensions of language use seems to me to be defective in another respect. There is one class of sentences which is strikingly absent: namely, those which purport to describe not an actual or a desirable state of affairs, but a possible state of affairs. That is, no provision exists for modality in the framework as described. In a forthcoming article (Wilson & Sperber, 1988a) the authors lump together “actual, potential or possible” attributed thoughts. It is not obvious that the same identification is appropriate in the characterisation of states of affairs, if only because of the problem of adequately defining the truth-conditional differences among the three. However, there seems to be no reason in principle why provision should not be made for a third branch on the bottom right of the diagram to account precisely for such a category. The main natural language realisation of such a category would then be conditional sentences of the kind: “If giraffes had wings, they would lay eggs” (p. 229). This is characterised as “interpretive” (*ibid*); but this classification seems no more appropriate than the characterisation of desirable states of affairs. The existence of a third branch would of



course suggest that there should be a congeneric equivalent on the interpretive side; and indeed in Smith and Smith (1988) it is essentially suggested that counterfactual conditionals (of the kind "If giraffes hadn't had wings, they wouldn't have been able to fly so well") filled precisely this alternative box. I now think that this conjecture (in line with remarks made earlier by S&W) was misguided, despite its initial appeal and the syntactic evidence available, and that conditionals, both indicative and counterfactual, should be characterised as descriptive, with the difference between them residing partly in the explicature of falseness associated with counterfactuals and partly in the degree of commitment to the plausibility of the particular conditional that the speaker is prepared to make, especially when this scale of commitment is morphosyntactically marked.

I think that the contrast between descriptive and interpretive use is one of the more important and insightful ones that S&W have come up with. I also think that there is room both for modification and extension along the lines indicated to provide an account which is more consistent, more transparent, and more comprehensive.

## Relevance must be to someone

Yorick Wilks

Rio Grande Research Corridor, Computing Research Laboratory, New Mexico State University, Las Cruces, N. Mex. 88003

From an artificial intelligence perspective, Sperber & Wilson's (S&W's) *Relevance* deals with many familiar issues in an odd way, particularly in that S&W remain determinedly unfamiliar with information-processing literature, even though they use the phrase and related ones a great deal. Because the précis of the book leaves out all their more precise claims and examples, I shall need to quote from the book directly. My general point, that their claims about processing are false, cannot be made, along with the quotation it requires, in the space available, but can be found in Wilks and Cunningham (1986). Here I shall restrict myself to the discussion of one example and the principle S&W use it to illustrate, arguing that no account of relevance in terms of beliefs and their consequences can ignore the need for processing models of the beliefs of others – or, as my title puts it, relevance is always to someone. This is a truism in artificial intelligence, at least since the pioneering work of Perrault and Allen (1980), but S&W's work remains within the process-free linguistic paradigm where one can consider the consequences of beliefs for communication without considering whose they are.

We shall refer to the excerpt from *Relevance* that follows as the mad-passerby example. One piece of cultural knowledge may prove necessary for understanding it: A flag-seller is a volunteer for a charity who stands on the pavement/sidewalk asking passersby to donate money; they receive a small paper flag in return.

(4)

Flag-seller: Would you like to buy a flag for the Royal National Lifeboat Institution?

Passer-by: No thanks, I always spend my holidays with my sister in Birmingham.

To see the relevance of the passer-by's response, the hearer must be able to supply something like the premises in (5), and derive something like the contextual implication in (6):

(5)

- (a) Birmingham is inland.
- (b) The Royal National Lifeboat Institution is a charity.
- (c) Buying a flag is one way of subscribing to a charity.
- (d) Someone who spends his holidays inland has no need of the services of the Royal National Lifeboat Institution.

(e) Someone who has no need of the services of a charity cannot be expected to subscribe to that charity.

(6)

The passer-by cannot be expected to subscribe to the Royal National Lifeboat Institution.

What is interesting about the passerby's reply is the very close connection that exists between seeing its relevance (or, more precisely, the relevance the speaker intended it to have) and being able to derive some contextual implication from it. It seems clear that someone who is unable to supply something like the context in (5) and derive something like the contextual implication in (6) will be unable to see the intended relevance of this reply. . . . (p. 121)

So far, the reader should be in a not wholly unfamiliar world: A restatement of the notion of enthymeme, or incomplete argument, well known to such individuals as Bochenski (1947) of the Polish School, who argued that understanding is a matter of finding propositions that make what a speaker says true; and Peirce's (1955) "abduction," the inference scheme in which the premise is established by the truth of the conclusions.

One might reply initially to the above excerpt, not as a fundamental criticism but as a form of mind-clearing, that what is interesting about it is the degree to which matters need not be at all as S&W describe them. There are two related issues here: (i) the point-of-view problem, or "Whose inferences are the ones characterised above?" and (ii) the alternative-hypotheses problem.

As to (i), the passage strongly suggests that in the above account it is a model of the hearer that is being offered ("the relevance the speaker [i.e., the passerby] intended it to have"), and that the information processing or inferencing is being done by the hearer. The additional premises in (5) are then those provided by the hearer, the ones he believes the speaker may believe (and, in something like the classic Gricean manner, believes that the speaker intends him to attribute to the speaker in the inference process). So, according to that view, the model proposed (to use AI-psychology terminology without apology) is of the hearer's model of the speaker.

As to (ii), it is clear that the hearer may in fact attribute a set of beliefs to the speaker quite different from those in (5), yet still derive (6), and hence "see the relevance" of the mad passerby's remark. Thus the final sentence in the quotation above is quite false.

An alternative set would be (5'):

(5')

- a. The Royal National Lifeboat Institution is a charity that provides cheap holidays for poor elderly people (cf. the lexical with the real semantics of "The Salvation Army").
- b. The speaker is a shabby elderly-looking person.
- c. Someone who already has holiday provision will not need the services of a charity providing it.
- d. Someone who has no need of the services of a charity cannot be expected to subscribe to that charity (the same as 5e).

It is of no importance that (5'a) is a false belief, as was (5e), for belief attribution in communication cannot require that we attribute to others only beliefs we happen to hold. If that were needed, communication would rapidly collapse, and we could not talk to those of political or linguistic beliefs opposed to our own. (5'b) has a special quality, in that it appears to be a belief of the hearer *about* the speaker, rather than about the speaker's own beliefs. This proposition might then be an inappropriate inferential construction by the hearer, if he believed the speaker was unaware of his own appearance.

Decisions like this, as to what "logical space" or "mental environment" the inferences are to take place in (e.g., within the hearer's view of the speaker, the speaker's view of the hearer's view of the speaker, or . . .) are, in our view, vitally important, though they have no place in S&W's scheme. They are the basis of artificial intelligence work in this field (e.g., Perrault & Cohen, mentioned above, or Wilks & Ballim 1987). No account that uses the term "processes" can be taken se-

riously unless it actually deals in psychological or computational processes and S&W do neither.

The issue here is simply the wrongness of S&W's "someone who is unable to supply . . . will be unable to see the intended relevance of the reply" quoted above. If that really means derive by means of (5), then it is just wrong, as we have shown. The difference here is much more than the lack of a footnote from S&W suggesting the possibility of other derivations: It is that they do not really accept that such inferences must be some particular individual's inferences, and so they feel free to opt for an "objective" set of hypothesised premises, ones they believe to be correct.

A moment's more ingenuity will provide yet further interpretations, which may have more contextual implications than the use of (5) or (5'):

(5')

a. People who spend holidays with relatives normally murder them sooner or later.

b. The Royal National Lifeboat Institution is a charity (same as 5b).

c. People who do not subscribe to charities also murder their relatives.

The sequence abc yields not only (6) but (6'):

(6')

The passerby is likely to murder or has already murdered his sister.

This last example is significant as well as baroque because it refutes S&W's most important and precise claim that, "other things being equal, an assumption with greater contextual effects [i.e. a greater number of non-trivial consequences] is more relevant; and, other things being equal, an assumption requiring a smaller processing effort is more relevant" (*Relevance*, p. 125). Again, since (5') appears on the face of it less complex and in need of processing effort than (5), though allowing the same consequence (6), it refutes the second part of the quoted claim.

---

## Editorial commentary

As a criterion for relevance, counting (nontrivial, contextual) implications seems to entail an infinite regress; for what then is to serve as the criterion for *relevant* implications? Nor is it clear how capacity considerations, likewise ungrounded in themselves, can halt the regress. Or are we, like St. Exupéry's star-counting businessman, just indiscriminate implication-collectors, restrained only by our limited resources?

---

## Authors' Response

### Presumptions of relevance

Dan Sperber<sup>a</sup> and Deirdre Wilson<sup>b</sup>

C.N.R.S., 15 Quai Anatole France, 75007 Paris, France<sup>a</sup> and Department of Linguistics, University College London, London WC1E 6BT, England<sup>b</sup>

Whether they approve of or attack *Relevance*, the accompanying commentaries illustrate its theses. They show that communication takes place at a risk, that mutual knowledge does not exist, that what is relevant to one differs from what is relevant to another, and that we are all more easily convinced of the relevance of what we have to say than of what we have to hear. As relevance theory

predicts, each of these commentaries communicates a presumption of its own relevance, and we have found them all relevant, though not always in the way their authors intended. We are grateful to all of our commentators and hope that our response will confirm the presumption of relevance which, however much modesty and wisdom tell us we should not, relevance theory tells us we cannot help but communicate. Our discussion will roughly follow the plan of the book and of its précis.

## 1. Communication

**1.1. Coding and inference.** One common worry with the Gricean account of communication is that it attributes to the hearer an implausible amount of reasoning about the speaker's intentions. We share that worry. This is one of the reasons we have tried to show that a number of issues (e.g., the extent of mutual knowledge and the intentions behind violations of maxims), resolved in the Gricean framework by reasoning, do not arise at all, and that the remaining issues can, in the case of successful communication, be resolved by a form of unconscious inference simpler than conscious reasoning.

Both Millikan and Pettit feel that too much inference about the speaker's intentions may still be involved in our framework. They particularly object to the view that an essential reason for a hearer to believe what the speaker told him is his recognition that she intended him to believe it, combined with his confidence in her. These commentators would rather have successful communication convey an assumption "immediately" (Pettit) or produce a direct "belief transfer" (Millikan).

A possible case of "belief transfer" among humans is the spreading of panic in a crowd, but most human communication, especially verbal communication, is not like this at all. Rather, it is achieved by means of ostensive stimuli, utterances in particular, which require attention to be understood, and which only partially encode their interpretations, if they do so at all. How is decoding to be supplemented? Millikan and Pettit don't discuss this question, and the only answer we can think of is that it is supplemented by inference. Inference goes from evidence to conclusions warranted by the evidence. As far as we can see, an utterance such as "he is coming" provides no evidence for the conclusion that William is coming unless it first provides evidence that the speaker intends it to be so understood. In addition, the reliability of the speaker must be taken into account.

But then, to say that understanding consists in forming a hypothesis about the communicator's informative intention, and that the fixation of beliefs derived from communication must be sensitive to considerations of the speaker's reliability, is not equivalent to saying that there are steps in the inferential process where the communicator's intentions and reliability are scrutinized. According to our account, the first interpretation of an utterance to be accessed and found relevant enough is taken to be the one intended by the speaker, unless there are salient reasons to think otherwise. Thus Pettit is right (in most cases) to point out that the relevant "features of context invoked in such inferences do not concern the speaker's intentions"; nevertheless, the conclusion of the inference is *about* these intentions.

Similarly, in many ordinary cases of communication,

the reliability of the speaker is taken for granted unless there are salient reasons to doubt it. We agree, then, with Davies when he argues (against McDowell 1980) that beliefs are sensitive to reasons and suggests that what this implies is “a condition about the absence of certain undermining doubts rather than about the presence of certain justifying beliefs.” To put it in other terms (and to show Levinson that we do not ignore “default inferencing”), in our model the speaker’s success in predicting what will be optimally relevant to the hearer, as well as her reliability, is assumed by default. This simplification of the inferential task involved in comprehension should help bridge the gap between Millikan’s and Pettit’s point of view and ours.

Millikan sees it as implausible that “extremely young children represent to themselves that the speaker *intends* that they should *believe* that *p*.” There seems to be some evidence, however, that children might develop such an ability in the course of their second year (Bretherton et al. 1981; Leslie, 1987). Agreeing with Leslie that “the relationship between metarepresentational capacity and communication in infancy remains an interesting but still open question,” we suggest that relevance theory may help to clarify it by providing a simpler account of the role of metarepresentation in communication. In particular, one might assume that the default values mentioned in the preceding paragraph are fixed values in the first stages of verbal communication.

Pettit makes a further point about the distinction we draw between (peripheral) decoding processes and (central) inferential processes. Citing Cummins (1983), he comments that since a process may be interpretable as inferential even though it does not explicitly represent inferential steps, decoding would seem to count as “a clear instance of inference.” But Cummins points out (1983, p. 53, n. 1) that not all computations over sentential representations are interpretable as inferential: “A device instantiating a transformational grammar transforms sentential inputs into sentential outputs, but the inputs and outputs are not related via any pattern of inference.” We assume (*Relevance*, pp. 71–72) that the mind has a variety of systems of nonconceptual representation (e.g., phonetic representation) and of noninferential computation (e.g., syntactic transformation). Because the input and output of a syntactic transformation are not related as premise and conclusion in an argument, the computation is not interpretable as inferential.

It is true, as Pettit points out, that linguistic input processes can be characterised as inferential in a broader sense: From the premise that Mary uttered sounds with a certain phonetic representation, Peter may be able to infer that she said she was leaving. But this does not affect the point at issue: The linguistic decoding processes embedded in this broader inferential process are not themselves inferential, in the sense that they are not truth-preserving (or strength-preserving) operations on conceptual representations, whereas the operations performed by our deductive device are inferential in just this sense.

**1.2. Mutual knowledge.** In *Relevance* we argue that the existence of mutual knowledge as defined by Schiffer (1972) is psychologically implausible, and that its existence has been assumed not on the basis of any kind of

direct evidence, but because it was needed to patch the code model of communication or to produce a failsafe inferential model. Once we aim, realistically, at describing communication in terms of an inferential heuristic rather than a code or an inferential algorithm, we need no longer burden ourselves with mutual knowledge; the weaker and psychologically more plausible notion of a mutual cognitive environment, or of mutual manifestness, suffices.

Bach & Harnish point out both that their approach (1979) is inferential and that they are very aware of the fact that communication is not failsafe. Their reason for adopting a mutual knowledge framework is merely “to limit the speaker’s intention and the hearer’s inference to information that each can expect the other to rely on.” On this, as on the analysis of speech acts, these authors are right to assert that their approach does not fall squarely under the criticisms we level at current pragmatics in general. This being so, they should unhesitatingly adopt mutual manifestness: It will do the job they want done as effectively as any realistic approximation of mutual knowledge, and it is much more economical.

Gibbs and Russell believe that our criticisms of mutual knowledge are in part based on our failure to realize that knowledge need not be consciously or explicitly represented. Gibbs suggests further that our distinction between knowledge and manifestness is merely a distinction between conscious and tacit knowledge, and suspects us of “‘sneaking’ mutual knowledge in the backdoor of [our] theory.” As we discuss and illustrate our theory (p. 40) we point out that humans can be said to believe tacitly, or virtually, what they are capable of inferring *demonstratively* from their mentally represented beliefs. The problem with mutual knowledge is not just that humans are incapable of having an infinity of beliefs explicitly represented in their mind. It is also that the infinitely many beliefs which together make up mutual knowledge are not demonstratively inferable from a finite set of premises. Hence they cannot even be held as tacit or virtual beliefs.

What humans are merely capable of inferring non-demonstratively from their explicitly represented beliefs is not a set of further “tacit” beliefs but a set of assumptions manifest to them. Hence “manifest” is, as we show at length in Chapter 1, Section 8 of our book, not at all equivalent to “believed,” let alone “known”: There is nothing surreptitious about our use of “mutual manifestness.”

Several commentators acknowledge the psychological implausibility of mutual knowledge but argue that it can be remedied by redefining the notion. Russell favors Halpern and Moses’s (1985) “definition” of mutual knowledge of proposition *p* as, in Russell’s informal paraphrase: “Everyone knows *P* and everyone knows it is mutually known.” With the definiendum repeated in the definiens, however, this purported definition seems to be blatantly circular and to provide no psychological enlightenment whatsoever.

McCawley argues that one should take mutual knowledge to be “knowledge on the part of a group of two or more persons as a whole and . . . not attempt to reduce such knowledge to knowledge on the part of each of those persons individually.” Yet it would seem that either you reduce mutual knowledge so understood to some set of



mental and environmental states or processes – but McCawley does not even suggest how this should be done – or else you commit yourself to ontological pluralism. Relevance theory, like most cognitive psychology, remains within a materialist, more specifically a token-physicalist, framework.

Hinkelman misdescribes mutual manifestness; she claims that it does not prevent regress (we agree that it does not prevent it, it merely makes it unnecessary) and she states, without argument but with a reference to a forthcoming article of Perrault (1987), that “default formulations of mutual belief are more satisfactory.” Davies makes an explicit proposal in this spirit: Mutual knowledge might be replaced by “mutual absence of doubt.” The general idea is that communicators do not need reasons to believe in the existence of a common code and of mutual knowledge, which will simply be taken for granted in the absence of reasons to doubt them.

The arguments against the standard formulation of mutual knowledge presented in *Relevance* carry over, however, to the default formulation: If some arguments cast doubt on the very possibility of mutual knowledge, then, for every proposition, there are reasons to doubt that it is mutually known by any two people. Of course, not everybody will be sensitive to these reasons, but this reduces the plausibility of mutual knowledge still further. If a simpleton never doubts mutuality, if some philosophically sophisticated people always doubt it, and if most people doubt it to varying extents, then no two people are ever likely to match in their absence of doubt (which makes systematic doubters right). Default formulations of mutual knowledge may be attractive qua formulations, but they come nowhere near explaining how true mutuality might actually be achieved.

We would like to ask Davies in what sense an absence of doubt may be *mutual*, as opposed to simply *reciprocal*: Communicating bees can be credited with a reciprocal absence of doubt (since they have no doubts of any kind), but clearly Davies wants to credit communicating humans with something stronger. We are not quite sure what, and why.

Following Lewis (1969) and Clark and Marshall (1981), Gerrig sees the solution to the psychological problems raised by the notion of mutual knowledge in the following induction schema. A and B mutually know that *p* if and only if some state of affairs *G* holds such that:

- (i) A and B have reason to believe that *G* holds.
- (ii) *G* indicates to A and B that each has reason to believe that *G* holds.
- (iii) *G* indicates to A and B that *p*.

Clark and Marshall (1981) give psychologically realistic examples of how this schema might work, and McCawley lists “unproblematic sources of mutual knowledge” in the same vein. Isn’t mutual knowledge so understood psychologically plausible after all?

The examples given by Clark and Marshall (1981) and by McCawley which satisfy the above induction schema are indeed realistic, but they are not examples of mutual knowledge, nor even of mutual belief. Peter and Mary are at home watching television. Does Mary believe that there is no general strike at the electric company? There is no way of telling. Yet Mary (like Peter) has reason to

believe that there is a state of affairs *G*, namely the fact that she and Peter are watching an electrically powered television set functioning normally, which indicates that *p*: namely, that there is no general strike at the electric company. Furthermore, *G* indicates to Peter and Mary that each has reason to believe that *G* holds. So, by the above schema, we should infer that *p* is a mutual belief held by Peter and Mary, when there is not even enough evidence to assume that either of them believes that *p*! The missing piece of evidence is this: Did Mary ask herself whether there was a general strike at the electric company? If she did, then she probably now believes that *p*. Otherwise she does not believe either that *p* or that *not-p*.

As we argued (p. 45 and p. 258, n. 29), what is taken by Lewis (1969), Clark and Marshall (1981), and Gerrig as a necessary and sufficient condition for mutual knowledge is merely a mutual cognitive environment. A mutual cognitive environment gives each of those who share it evidence of the other’s beliefs, including, to some extent, beliefs about the other’s beliefs, provided they also have evidence of where the other’s attention has been and is going. (One way of knowing that is to deliberately manipulate the other’s attention by means of ostensive behaviour.) However, a mutual cognitive environment does not provide sufficient evidence for mutual knowledge or belief; nor, as far as we can see, does anything else.

What Gerrig, Gibbs, McCawley and others give as examples of mutual knowledge are good examples of mutual manifestness. McCawley even uses one of our own examples and writes that since the information alluded to is *manifest*, it “will count as mutual knowledge!” So we are all referring to the same kinds of situation. But the difference is not merely one of terminology, it is primarily a difference of analysis, and it leads to different explanations of communication.

Gerrig, Gibbs, and McCawley maintain not only that mutual knowledge is possible, but also that it is necessary. What they do, basically, is to evoke in lesser (McCawley), or greater (Gerrig, Gibbs), detail the standard mutual knowledge analyses of reference assignment and presupposition (McCawley), implicature and irony (Gibbs), and accidental irrelevance (Gerrig) as if this were enough to establish that no alternative analysis is possible. Note that our objection to analyses which presuppose mutual knowledge is not that they do not work: It is that their presupposition is false. If, as we claim, mutual manifestness permits enough coordination for successful communication to take place, so, a fortiori, would mutual knowledge, if it existed. Gibbs and McCawley do not discuss our approach to reference assignment, presupposition, implicature and irony (chap. 4, sec. 3, 4, 5, 9). We read their remarks as a respectable expression of theoretical conservatism, but not as arguments that we must answer. Gerrig also ignores our analysis of accidental irrelevance (pp. 159–60); but at least he tries to show that relevance theory without mutual knowledge would get the wrong results. In fact, as we will show below (sect. 3.3), what he gets wrong is relevance theory.

If mutual knowledge does not exist, how does the communicator anticipate, as she should, the thought processes of the audience – a question raised in different terms by Morgan & Green, Bach & Harnish, and Millikan? The answer lies in the fact that it is not only in

communicating but in every form of interaction that humans have to anticipate each other's thoughts. We claim that the case of the communicator does not differ essentially from others. We have no novel and elaborate account of how people can know each other's thoughts in advance of communication. We believe, however, that relevance theory can help to develop such an account. If human thought processes have not only a variety of changing purposes but also a single constant aim – that of maximizing relevance – then they are easier to predict.

Moreover, since accessibility affects relevance and is itself affected by immediately preceding thoughts, predictions are much easier to make (and to study) in a situation of ongoing communication, where the communicator knows what the audience has just been attending to. Note further that, according to our account, the communicator's anticipations of the audience's thought processes need not be accurate. Communication will still succeed if the audience is capable of recognizing the communicator's mistaken anticipations.

**1.3. Gricean themes.** In Chapter 1, we discuss Grice's views (see references in *Relevance*) on meaning and communication and try to show that they contain fundamental insights but also serious shortcomings. In Chapters 3 and 4, we compare Grice's views and our own. Some commentators (**Bach & Harnish, Clark**) feel that we depart too much from Grice's approach; others feel that we have departed from it much less than we claim (see sect. 3.4 below).

Grice argues that what we call an informative intention must be "overt" and we agree, but we find the various analyses of that notion in the Gricean literature unsatisfactory. **Bach & Harnish** rightly point out that we do not discuss the rather vague suggestion of Grice (1969) that overtness might be characterized by the absence of covert intentions, a suggestion criticized by Schiffer (1972). They also maintain that, contrary to what we say, a reflexive intention (a communicative intention *I* which includes the subintention that the audience recognize *I*) need not be expanded into an infinitely long formula. Our claim (pp. 256–57), however, is not that the expansion is formally necessary, merely that without it the content of the intention is beyond the individual's grasp; since with expansion the content of the intention is even more obviously beyond the individual's grasp, we conclude that this is not the right way to analyse overtness. The right way, we suggest, is to say that an informative intention is overt if and only if it is mutually manifest.

Another problem with Gricean accounts is that they explain the intuitively correct interpretation of an utterance by showing that it satisfies certain criteria which, for all we know, very different, intuitively incorrect interpretations might also satisfy. **Bach & Harnish** say that we "do not show that Gricean accounts allow this possibility when communication *succeeds*." But surely the burden of proof is on the Griceans. In our theory, we show that an utterance cannot have more than one interpretation consistent with the principle of relevance. If we did not, would not **Bach & Harnish** object that our theory might fail to exclude intuitively incorrect interpretations?

**Clark** criticizes us for not "embracing Grice's and Austin's full insight" that "communication must be

viewed as part of a theory of action." Our characterization of relevance, he complains, "is always divorced from what the participants in a discourse are really doing." Our position is "like claiming that the only purpose I need in stepping on the car's accelerator is to put more gasoline into the carburetor."

Still, **Clark** notes that Grice himself did not do much to develop his "insight"; and work on speech acts in the Austinian tradition, however rich, could hardly be described as particularly "insightful" (chap. 4, sec. 10). Moreover, it is far from clear that the very idea of a "theory of action" has much to commend it: Is action, or human action, a natural kind to be studied as an isolated whole? (For similar doubts about a theory of communication in general, see pp. 2–3).

Generally, of course, communication is not an end in itself but a means to an end. So is driving a car. But what makes communication and car driving so useful is that, with essentially the same procedures, an indefinite variety of possible ends can be pursued. A technological study of car driving might end by relating different driving styles to different driving purposes – taking the family on a picnic or car racing – but it certainly should not start there. Similarly, our aim is to describe the mechanism of ostensive–inferential communication in general terms, and therefore to develop a notion of relevance common to all cognitive and communicative endeavours. Contrary to what **Clark** suggests, there is nothing intrinsically paradoxical about such an aim, nor is the alternative approach he favours obviously superior.

**1.4. Ostensive communication.** We were surprised that, whereas our proposal to replace mutual knowledge by mutual manifestness was hotly debated by six commentators, only **Bach & Harnish** had remarks to make about what we feel is a more important proposal, to replace the Grice/Strawson definition of meaning and communication by our characterization of ostensive–inferential communication.

We are grateful to **Bach & Harnish** for noticing an inconsistency in our formulation: In introducing our definition of a communicative intention (pp. 60–61) we incidentally suggest that ostensive communication has to be intentional, when in fact we believe, and state explicitly after defining ostensive communication (pp. 63–64) that it need not be – although unintentional ostensive communication is a very rare occurrence. We answer another question put by **Bach & Harnish**, "How does one identify a stimulus as an instance of ostensive communication?" at length in Chapter 3, Section 6.

**Bach & Harnish's** third point about ostensive communication is of substantive interest: They note that our definition of ostensive communication "does not require uptake, since it specifies only that the speaker's informative intention<sup>1</sup> . . . be made mutually manifest, not that the hearer actually recognize this intention. For S&W an assumption is manifest to someone merely because he is *capable* of representing it and of accepting its representation as true (p. 39); but then the thought of it might never occur to him and might never play a role in his thinking." **Bach & Harnish** seem to present this consequence of our definition of communication as an objection; we believe on the contrary that it is one of the ways in which our definition improves on previous ones.

We will illustrate this with two examples, the first intended merely to show that the association between uptake and communicative success is not self-evident, the second to show that there is a positive analytical benefit in dissociating the two.

Consider a botanical garden where every plant on exhibition has a label indicating its name. Each label is an ostensive stimulus, or an utterance in Grice's sense, and its display constitutes an act of communication. Now if communication implies uptake, then those labels which happen not to have been read by any visitors must be regarded as cases of communicative failure. If manifestness, rather than uptake, is the condition for communicative success, then only those labels which are invisible or illegible count as communication failures; all the labels which modify the cognitive environment of the visitors by making them *capable* of recognizing that the curators of the garden intended to inform them that this or that plant is called so-and-so are cases of communicative success. This example should show at least that it is not intuitively obvious in every case that communicative success implies uptake.

Now take a more ordinary example of verbal communication, one we used both in our book (pp. 193–99) and in our accompanying *BBS* précis:

- (1) (a) *Peter*: Would you drive a Mercedes?  
(b) *Mary*: I wouldn't drive ANY expensive car.

We argue that (1b) has not only the strong implicature (2) but also a range of weaker implicatures such as (3):

- (2) Mary wouldn't drive a Mercedes.  
(3) Mary wouldn't drive a Jaguar.

Weak implicatures such as (3) need not achieve uptake in order to be communicated: On hearing Mary's reply, Peter might not think of (3) at all. Even so, Peter might, on a later occasion, wonder whether Mary would drive a Jaguar and then realize that she has already communicated to him that she would not.

A definition of communication according to which the addressee must mentally represent every assumption communicated to him is incompatible with the notion of weak implicature illustrated by (3), as well as with the explanation of poetic and other hazy effects of communication that we build around this notion. We see this as a strong reason for not making uptake a necessary condition for communicative success. This is not to deny, however, that uptake plays an essential role in ordinary communication.

Take the strong implicature (2) in the above example: If it becomes manifest at all to Peter, he is sure to entertain it since it provides him with the very information he asked for. (Moreover, none of the implicatures of (1b) will become manifest to Peter if he does not first entertain the explicit content of the utterance.) In general, for a set *I* of assumptions to be communicated, some members of the set have to be actually entertained; and in general there are few acts of communication without uptake. However, this is so for practical rather than for definitional reasons; moreover, it does not imply that every assumption communicated in an act of communication need achieve, or is even likely to achieve, individual uptake.

## 2. Inference

A crucial element in our account of communication and cognition is the operation of a deductive device which can automatically compute the full set of contextual implications of a given assumption in a given context. Here are some objections and questions.

Russell objects that this method of operation is "absurdly impractical" and that some control method must be used to ensure that only "useful" inference paths are explored. This misses one of the major points of the book. A definition of relevance which itself presupposed a notion of cognitive "usefulness" would be profoundly unhelpful. On the other hand, given an adequate concept of relevance, a notion of cognitive "usefulness" can be defined. For example, a "useful" path in inference is one that yields enough contextual effects to be worth the effort.

In our framework, inference is "controlled" by two main methods, only one of them internal to the deductive device. In the first place, the distinction between trivial and nontrivial implications imposes a substantial internal constraint on the class of implications the device is capable of computing, but one that is definable independently of such notions as cognitive usefulness.<sup>2</sup> In the second place, the class of inferences drawn is determined by the set of assumptions submitted to the deductive device. The choice of assumptions is determined by such factors as context selection, perceptual salience, and focus of attention. Inferences tend to follow "useful" paths because, we argue, all these factors are geared to the maximization of relevance.

Russell, pointing out that in our framework all inference rules are explicitly represented, objects that such a scheme cannot succeed, because "whatever combines the inference rules with their premises must also be an inference rule, but procedurally represented." But inference rules are not part of the data base in our framework, and they are not in propositional form. We see the distinction between representation and computation as fundamental, and, as is implicit in the (transformational) formalism and we explicitly state (pp. 84, 89), our inference rules fall on the computational rather than the representational side. So, strictly speaking, our inference rules don't "combine" with premises, they operate on them; they are procedures in an ordinary sense. How the rules operate on the premises is not determined by further rules (and in what sense would they be *inference* rules anyhow?) or representations of any kind, but by the way the deductive device is constructed, as we suggest (pp. 94–95).

Russell objects that a deductive device with access only to elimination rules would be unable to use "commutativity laws and axioms defining inductive domains." But because axioms in our framework are treated as information in encyclopedic entries rather than as inference rules, the deductive device has access to them as premises. However, Russell has drawn attention to a genuine, though easily solvable, problem as far as commutative rules are concerned.

There is indeed no a priori reason why the deductive device should not have access to such rules, or, more generally, to analytic rules which reorganise the concep-



tual structure of assumptions without eliminating any of their logical content. In Sperber & Wilson 1982 (pp. 102, 109), we cite empirical evidence for two such rules, one converting assumptions of the form  $(P \vee Q) \rightarrow R$  to conclusions of the form  $(P \rightarrow Q) \& (P \rightarrow R)$ , the other converting assumptions of the form  $(P \& Q) \rightarrow R$  to conclusions of the form  $P \rightarrow (Q \rightarrow R)$ . These are allowed for in our definition of elimination rule (p. 86, *Relevance*). But we might also want a rule converting an assumption such as: *X is the cousin of Y* into a conclusion such as: *Y is the cousin of X*. Such a rule would fall into a gap between our definitions of elimination rule and introduction rule (p. 96). We might then either redefine an elimination rule as anything that is not an introduction rule, or allow the deductive device to use rearrangement rules.

Seuren sees it as an awkward consequence of our system that it rules out the inference from (4) to (5):

- (4) Henry and Jack came in.
- (5) Two people came in.

Such examples are dealt with in Sperber & Wilson (1982, p. 102). The move from (4) to (5) involves the ability to count (more generally, the ability to perform mathematical calculations). We see no reason to attribute such abilities to the deductive device. Nor do we need to. Anyone who can count may, on the appropriate occasion, construct the additional premise "If Henry and Jack came in, then two people came in," which will combine with (4) to yield (5) as a contextual implication.

Macnamara has a more general objection to the ban on introduction rules. "How," he asks, "could there be anything for elimination rules to work on, if there were no introduction rules?" The answer is that elimination rules work on any set of assumptions submitted to the deductive device. As pointed out in *Relevance* (pp. 81–83), unconscious deduction is not the only source of factual assumptions: Perception, linguistic decoding, assumptions and assumption schemas stored in memory, and conscious reasoning are further sources not affected by the ban on introduction rules.

The deductive device operates on conceptual representations in virtue of their logical structure and the concepts they contain. We define a concept as an address in memory which gives access to three types of entries – logical, encyclopedic, and lexical. Unlike semanticists who see the semantic properties of a word as specified in its lexical entry, we see them as provided by the logical entry filed at the same address. As the context should make clear, this is what we mean when we say (p. 90) that "the 'meaning' of a word is provided by the associated concept." Macnamara, understanding "provided by" in the sense of "consists of," takes a concept to be all the information filed under a concept and goes on to establish that "things are getting seriously muddled."

Things are indeed muddled. The semantic theory Macnamara attributes to us is strange: The syntactic and phonological properties of a word are considered part of its meaning, "giraffes live in Africa" comes out as analytic, and so on. This directly contradicts the views we outline in the section from which this one misunderstood sentence is taken: In particular, we stress the parallel between the classical analytic/synthetic distinction and our

logical/encyclopedic distinction (p. 88). We even discuss the logical entry for "giraffe" and suggest that it contains not more, but much less, than a set of necessary and sufficient conditions for something to fall under the concept.

Macnamara comments that we proceed as though we had "no obligation to account for the existence of logic of the type one finds in the textbooks of formal logic." It is true that we do not feel obliged to account for the existence of formal logic: This is not our topic. But we believe that our approach does throw some light on the issue. In *Relevance*, we concentrate on spontaneous inference. As we indicate (p. 75), "Spontaneous inference plays a role even in scholarly interpretation, whereas scholarly thinking is a rather exceptional human endeavour, even for scholars. The study of spontaneous inference is thus a necessary prerequisite to a proper investigation of all forms of human inference." More specifically, we show (p. 102), how the deductive device, because it monitors contradictions, and in spite of its incomplete generative capacity, yields intuitions of logical validity or invalidity for inferences it cannot by itself perform. These might well be the "basic intuitions upon which formal logic rests" and for which, argues Macnamara, "an account of everyday reasoning must at least give the psychological grounding."

Hinkelman argues that a deductive device with access only to elimination rules can still generate infinite sets of conclusions from a finite set of premises, contrary to what we claim. Citing the example "She is a citizen if her mother is a citizen," she comments that "Any logic with modus ponens and function terms can regress infinitely." The deductive device is not "a logic," however; it is a computational mechanism limited in its operations not only by the rules it applies, but also by the way it applies them. Hinkelman's example is an ill-chosen one in any case, because it does not even logically warrant any regress. To make her objection stand, she should have chosen a more appropriate formulation, say: "Any woman is a citizen if her mother is a citizen"; and she might have added to it a finite number of premises of her choice. Then, if she found elimination rules independently justified in our framework which started the deductive device on an infinite regress, she would have had a genuine objection.

In an earlier review of selected fragments from the first half of *Relevance*, Seuren accused us of failing to define *contextual effects* (Seuren 1988). Having had the definition drawn to his attention, he has now read and misunderstood it. Contextual effects are contextual implications, strengthenings, and contradictions; we provide a two-clause definition (p. 260, n. 26), the first clause characterizing strengthenings and contradictions, the second contextual implications. Seuren, conceding that the second clause defines contextual implications, maintains that the first clause does not define anything, and that the notion of contextual effect is therefore undefined.

The claim made by the first clause of our definition is as follows: Let two assumptions with the same content but different strengths count as two different assumptions. The contextualisation of P in C has no contextual effect if and only if: {Conclusions of C} is a subset of {Conclusions of P  $\cup$  C}. Seuren believes that this fails to define

anything, because “whatever follows from a set of propositions  $C$  also follows from  $C$  enriched with any set of propositions.” Not in our framework. Suppose that  $P$  contains the single assumption  $p$  and that  $\text{not-}p$  is a conclusion of  $C$ . When  $\{P \cup C\}$  is placed in the memory of the deductive device, the result will be a contradiction. Suppose, further, that  $p$  is stronger than  $\text{not-}p$ , which is, accordingly, automatically erased. Then  $\{\text{Conclusions of } C\}$  – which contains  $\text{not-}p$  – will not be a subset of  $\{\text{Conclusions of } P \cup C\}$ , which does not; and the contextualisation of  $P$  in  $C$  will be correctly characterized as having some contextual effect. Similar arguments apply to contextual effects achieved by strengthening, as readers may check for themselves.

### 3. Relevance

The notion of relevance put forward in *Relevance* is novel in two ways: It concerns relevance in a context or to an individual rather than to, say, a topic; and it takes into account not only the cognitive effects of the relevant information but also the effort required to achieve these effects.

**3.1. Defining relevance.** Some commentators (Clark, Levinson, McCawley, Morgan & Green) think our definition of relevance fails to do justice to pretheoretical intuitions. Utterances are relevant, they feel, to purposes, goals, topics, questions, interests, or matters in hand.

We define relevance in a context and to an individual. We say what a context is, how it is constructed and how, once constructed, it affects cognition and comprehension. One reason we did not set out to define relevance to a purpose, goal, and so on, is that we had no idea how to answer the analogous questions for any of these terms: We could not define them; we could not say how they were selected or constructed; and we could not say how, once selected, they affected comprehension.

Given a definition of relevance in a context, and a method of context construction, however, there is no reason that assumptions about the goals and purposes of the individual, or of the participants in a conversation, should not form part of the context and give rise to contextual effects in the usual way. Such assumptions are likely to be particularly rich in contextual effects, since purposes and goals imply plans for action. We see no incompatibility between our work and a belief in the importance of goals, purposes, and plans; on the contrary, relevance theory sheds light on how these important notions may play the roles they play.

McCawley is unhappy about our claim (pp. 216–17) that topic-relevance, if definable at all, is a derivative notion (like cognitive usefulness, discussed above, sect. 2, para. 2). Here is an illustration of what we mean. Bransford and McCarrell (1977, p. 391) report the results of an experiment designed to show the importance of topic in comprehension. Subjects read a passage which was incoherent at first sight; told that the topic was *washing clothes*, they readily understood it. From the standpoint of relevance theory, what has gone on is simply that the mention of washing clothes has provided access to a chunk of encyclopedic information about washing clothes, in the context of which the passage

becomes relevant, hence intelligible. Results which seem to provide decisive evidence about the role of topics provide equally decisive evidence about the importance of contexts. Furthermore, it is easy to think of passages which would remain incomprehensible even given a topic, for lack of encyclopedic information. Conversely, it would be interesting to see whether (as we predict) comprehension would be achieved if the encyclopedic background were made highly accessible but nothing amounting to a topic was suggested. We see topics as derivative in the sense that their role (if any) is merely to provide access to contexts, which play a decisive role in every act of comprehension.

Is relevance a “unitary” concept – that is, is it a useful theoretical concept – or does it merely impose an arbitrary link between effect and effort, as McCawley and Russell suggest? Today Peter has two goals: One is to rest, the other is to get a good suntan. He could rest inside or he could get a suntan jogging on the beach; so there would be little theoretical justification for linking the two goals under a single concept. Today Mary has two goals: One is to impress Aunt Zelda by inviting her to a good restaurant, the other is to spend as little as possible. Her two goals are linked and can usefully be described as one: getting the best value for money within a certain range. Generally speaking, when a cost is incurred for a benefit, or an effort for an effect, there is nothing arbitrary in linking them under a unitary concept of efficiency. Relevance, as we describe it, is of course a type of efficiency (see pp. 46–50).

**3.2. Assessing relevance.** Bach & Harnish, Clark, Levinson, Morgan & Green, and Russell ask: How is processing effort measured, and how is it balanced against contextual effects? In *Relevance* (pp. 78–81) we distinguish comparative from quantitative concepts and argue (pp. 129–32) that although it would be possible, and of some interest to computer scientists and AI specialists, to define a quantitative concept of relevance, the concept of interest to psychologists is comparative rather than quantitative: Relevance, as it affects cognition, is not computed or numerically measured but monitored or assessed, yielding only gross absolute judgments and, in certain types of cases only, finer relative judgments. Suppose that the brain is sensitive to the amount of reorganization brought about by the processing of some information and to the expenditure of energy thus incurred, just as it is sensitive to changes of posture and expenditure of energy in the case of bodily movement. This is very vague – hopelessly so, some AI people may think – but it is not so vague that it could not be false, and it is what we are claiming anyhow.

Some independently assessable factors affect relevance in regular ways. For example, lack of contextual effects entails lack of relevance. The stronger the assumptions, the greater the contextual effects. The more frequent and/or recent the use, the smaller the processing effort. Processing more information in the same context, or the same information in a larger or less accessible context, involves greater effort. Greater effects, for the same or less effort, entail greater relevance; less effort, for the same or greater effects, entails greater relevance. If we are right, these factors affect the working of the cognitive system both at a given moment and over time: They affect

the individual deciding how best to allocate his processing resources, the speaker deciding what to say and how to say it, the growth of heuristics governing context-selection and focusing of attention, the organisation of memory, and the evolution of peripheral and other automatic systems and processes.

Russell makes four comments on the role of effort in our theory: first, that our notion of effort is far too flexibly defined to be a theoretically adequate tool for a real pragmatic system; second, that the methods we propose for assessing effort are ad hoc and introduced only in order to rationalise a desired interpretation; third, that it seems unlikely such minuscule amounts of effort would play a significant role, except in highly time-pressured situations; and fourth, that to place all the burden on an effort principle, implemented in an ordering on context accessibility, “abdicates the responsibility for a theory, passing it on to those studying memory organisation, who are already overloaded with ordering constraints passed on by other theories such as the possible-worlds theory of counterfactuals.”

These comments are inconsistent. It does not seem wise, within the space of a single paragraph, to describe people both as handing over all the work to others and as doing it themselves (however badly). Moreover, to say that our proposals “rationalise a desired interpretation” is just another way of saying they work; the job of judging whether they are “ad hoc” should surely be left to the psychologists. And the effort principle (which is not, incidentally, restricted to context selection, even in the examples Russell cites) is consistent with all the experimental results we know of.

The results reported by Cutler, for example, provide excellent confirmation of our claims. A speaker aiming at optimal relevance must intend, on the one hand, to achieve adequate contextual effects, and on the other hand to put the hearer to no unjustifiable effort in obtaining them. As Cutler comments, distortion of word onsets disrupts word recognition far more than distortion of later segments, and speakers make phonological choices in such a way as to minimise disruption to the listener – or, in our terms, to achieve the intended effects as economically as possible. Compare our remarks (p. 213) that “greater disruption [of the intonation contour] implies greater processing effort and, other things being equal, lowered acceptability” and that “stress placement, like other stylistic features, should be looked at in terms of processing effort.” We give ample evidence (chap. 4, sect. 5 and 6) that variations – even minute variations – in linguistic form can have a profound effect on pragmatic interpretation. We argue that relevance theory can shed new light on many aspects of style.

THE EDITORIAL COMMENTARY makes the following objection: “As a criterion for relevance, counting (non-trivial, contextual) implications seems to entail an infinite regress; for what then is to serve as the criterion for *relevant* implications?”

Our reply is, first, that relevance in humans is assessed not by counting implications but by monitoring contextual effects (contextual implications being just one type) and processing effort. Second, the relevance of an assumption is a function of the context in which it is processed. Suppose  $p$  is processed in context  $C$ , and  $q$  is a contextual implication of  $p$  in  $C$ . Then the relevance of  $q$

in  $C$  is the relevance  $q$  would have had if it had been processed on its own in  $C$ . A more useful notion, the one the EDITORIAL COMMENTARY may have intended, is that of the contribution of  $q$  to the relevance of  $p$  in  $C$ . According to our definitions, each contextual implication of  $p$  in  $C$  contributes to its relevance, but there is a sense in which they do not contribute to it equally, not because of qualitative differences in their content, such as “interestingness” or “topicality” (which would make our definition of relevance useless), but because of the formal properties of the system we describe.

An assumption  $p$  has only a finite number of contextual implications in a finite context. Some of these contextual implications (most of them, in most cases) are derived from others: The nontrivial implications of a contextual implication of  $p$  are themselves contextual implications of  $p$  (unless they are deducible from  $p$  alone or  $C$  alone). So the contextual implications of  $p$  differ in the number of further contextual implications of  $p$  they themselves imply and for the deduction of which they must first have been deduced. More generally, contextual implications of  $p$  differ in the size of the fraction of contextual effects of  $p$  they determine. (Incidentally, this difference in the contribution of various contextual implications to relevance plays an obvious role in the identification of implicatures.)

The deductive device, lacking insight, computes all the implications of any set of premises submitted to it, without considering the further effects these implications might determine (subject to the optional constraints mentioned in chap. 2, n. 23).<sup>3</sup> It then stops when all the implications have been computed (and not, contrary to what the EDITORIAL COMMENTARY takes us to believe, because of capacity considerations). The effects achieved and the effort expended (including the effort needed to describe the stimulus and access the context) determine relevance. There is no regress. The deductive device is indeed, in the Editor’s phrase, an “indiscriminate implication-collector,” but we human beings are not: We choose the premises the implications of which the deductive device computes.

**3.3. Understanding the principle of relevance.** The backbone of *Relevance* is the idea that the goals of human cognition and communication are related but distinct. The link between communication and cognition is relevance. Cognitive processes are aimed at maximal relevance. In ostensive–inferential communication the relevance aimed at, or at least presumed to be aimed at, is not maximal but *optimal* relevance.

The argument goes as follows. First, there is a principle of relevance: that every ostensive stimulus communicates a presumption of its own optimal relevance. Second, an ostensive stimulus is optimally relevant to an addressee if and only if it has enough contextual effects to be worth his attention and puts him to no unjustified processing effort in accessing them. Third, a hypothesis about the communicator’s informative intention is consistent with the principle of relevance if and only if a rational communicator might have expected the ostensive stimulus, on that interpretation, to be optimally relevant to the addressee. Finally, the first hypothesis tested and found consistent with the principle of relevance is the only hypothesis consistent with the principle of relevance, and it is the



one the hearer will choose. The fact that every ostensive stimulus has at most one interpretation consistent with the principle of relevance makes the goal of communication, at least on some occasions, achievable.

It took us some time to disentangle these notions, and some commentators are, not surprisingly, confused.

The greatest confusion consists in taking us to be advocating a criterion of maximal relevance according to which the right interpretation of an ostensive stimulus is the most relevant one. Wilks, who succumbs to this confusion, chooses to discuss the real-life example of a callous (rather than “mad”) passerby which we use (pp. 121–22) merely to illustrate the point that relevance and contextual effects are linked. Wilks invents an extravagant interpretation of the passerby’s utterance – an interpretation which is therefore both highly relevant if true and totally implausible – to show that the most relevant interpretation need not be the right one (very true – we ourselves stress the point, pp. 158–60), as if this somehow confuted relevance theory. Generally speaking, we agree with Wilks that the views he attributes to us are badly mistaken; however he is mistaken in attributing them to us.<sup>4</sup>

One of Hinkelman’s criticisms seems to imply that she too attributes to us a criterion of maximal relevance. She objects that, since comparisons of relevance are possible only in some cases, it may be impossible to compare the relevance of different interpretations of an utterance and choose the most relevant one. Concerning the criterion of consistency with the principle of relevance, the question to ask of a candidate interpretation is not “Is this the most relevant interpretation?” but “Could a rational speaker have thought that this utterance, on this interpretation, would have enough contextual effects to be worth my attention, and would also put me to no unjustifiable effort in achieving them?” As we point out (pp. 165–71), comparisons of relevance are required on the first approach, but not on the second. The first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance. Moreover, contrary to what Hinkelman takes us to believe, we do not at all assume that the more easily computed interpretations are the more relevant ones (p. 168).

Like Wilks and Hinkelman, Seuren takes us to claim that the hearer should compare all possible interpretations of an utterance before selecting one. Unlike them, he claims to have textual support for this mistaken belief, citing a passage (p. 165) in which we discuss a case in which all possible hypotheses are compared before the best is chosen. We go on to point out (p. 166) that this method of hypothesis selection, though appropriate in some cognitive domains, is not appropriate for use in comprehension. In other words, the very passage in which we reject the view that the correct interpretation of an utterance is the most relevant one is cited by Seuren as textual support for attributing this view to us.

Another confusion consists in taking us to be advocating a criterion of consistency with the *presumption* (rather than the principle) of optimal relevance. Gerrig succumbs to this confusion when he objects that, on our account, accidental irrelevance would lead to misinterpretation. We ourselves consider a case in which Mary expresses proposition *p* intending to inform Peter of *p*,

without realizing that he already knows it (pp. 159–60). Though *p* is irrelevant to Peter, and therefore inconsistent with the presumption of relevance, Mary’s utterance is consistent with the principle of relevance: What this requires is that Mary might have *thought* it would be optimally relevant to Peter. Given their common cognitive environment, Peter has no difficulty whatsoever in attributing this thought to Mary and therefore recognizing the intended interpretation, and, *mutatis mutandis*, this holds for Gerrig’s example too.

Several commentators are confused about what is presumed to be relevant in the presumption of relevance. Thus Levinson asks whether the principle of relevance applies to surface structures, semantic representations, explicated logical forms, or what? The answer is that it applies to none of these, but to an ostensive stimulus, in particular to an utterance as a whole. The relevance of a stimulus is defined in Chapter 3, Section 6; the special case of utterances is discussed in Chapter 4, Section 2.

McCawley and Clark present objections based on the assumption that, for us, what is presumed relevant in the case of an utterance is the proposition it expresses. McCawley writes that “‘Nice day, isn’t it?’ comes out as ‘irrelevant’ in S&W’s scheme.” Clark (whose more specific concern with literary discourse we will discuss later) writes: “Many types of discourse have more than one distinct layer of action or communication (Bruce 1981; Clark 1987; Goffman 1974 [he could also have mentioned Sperber 1975]). S&W, however, presuppose that all communication is flat,” meaning that we take into account only the relevance of the proposition expressed.

There are cases (such as the one we discuss pp. 159–60, to which McCawley refers) where, if an utterance is relevant at all, it is through the relevance of the proposition it expresses; but we do discuss other possibilities:

(i) The speaker expresses proposition *p*, which the hearer manifestly knows already (pp. 247–49). The speaker’s utterance may nonetheless be optimally relevant to the hearer. An utterance, like any other phenomenon, makes manifest a variety of assumptions, any of which may contribute to its relevance. By expressing *p*, the speaker might make manifest to the hearer that she is aware of *p*, is prepared to admit it, is happy or sad about it, and so on.

(ii) Mary expresses proposition *p*, which is manifestly incompatible with Peter’s firmly held beliefs. Mary’s utterance may nonetheless be optimally relevant to Peter. It may make manifest to him that *she* believes *p*, or is prepared to maintain *p* despite him (p. 181).

(iii) Mary expresses proposition *p*, which is manifestly irrelevant to Peter. The fact that someone chooses to express an irrelevant proposition may itself be highly relevant: For example, it may make manifest a desire to change the subject or it may be used to illustrate an abstract point (p. 121). In the case of phatic communication, such as McCawley’s “Nice day, isn’t it?,” Mary may express an irrelevant, or weakly relevant, proposition, in order to make manifest a variety of assumptions concerning her attitude to Peter, her willingness to engage in conversation, and so on.

There is a further confusion between, on the one hand, the notion of optimal relevance (defined and discussed, pp. 158–71), which plays a fundamental role in our theory of ostensive communication, and, on the other hand, the

notion of optimal processing (defined, p. 144), which plays a minor role in defining relevance to an individual (pp. 145 and 152–53). A phenomenon makes manifest a variety of assumptions, each of which could be processed in a variety of contexts. A phenomenon is optimally processed when the best possible combination of assumption and context is chosen: that is, the combination that contributes most to overall relevance.

Optimal processing is presupposed in our definition of optimal relevance, since the relevance in question is that of a phenomenon (the ostensive stimulus) to an individual. However, the two are not at all equivalent, if only because a phenomenon can be optimally processed but irrelevant (here optimal processing merely involves filtering out the information), whereas optimal relevance implies enough relevance to be worth the individual's attention. We believe that Levinson, who talks of optimising relevance but describes the principle of relevance as a function of contextual effects balanced by processing cost, and C. Smith, who talks of the principle of relevance as seeking an optimal balance between effort and effect, have conflated optimal relevance with optimal processing.

Bach & Harnish ask: "How do speakers and hearers follow the principle of relevance?" They don't. As we insist (p. 162), the principle of relevance is a generalization, not a rule or a maxim. "Communicators and audience need no more know the principle of relevance to communicate than they need to know the principles of genetics in order to reproduce. Communicators do not 'follow' the principle of relevance; and they could not violate it even if they wanted to. The principle of relevance applies without exception: every act of ostensive communication communicates a presumption of relevance."

Quoting the very passage which answers Bach & Harnish's question, Morgan & Green want to qualify it by pointing out that "speaker and interpreter have to 'know' the Presumption of Optimal Relevance in order for their [S&W's] claims about how communication is accomplished (pp. 165–68) to be correct." Let us qualify Morgan & Green's qualification by stressing that what must be mutually manifest (rather than known) to communicator and audience is not the presumption of relevance in its general and abstract form, but merely the fact that a particular act of communication was intended by a particular communicator to appear optimally relevant to a particular audience.

**3.4. The principle of relevance and Grice's maxims.** Most modern pragmaticians accept some version of Grice's cooperative principle and maxims of truthfulness, informativeness, relevance, and clarity. Though deeply indebted to Grice, we do not follow him here. Neither the principle of relevance nor the presumption of relevance is a maxim addressed to speakers, known by hearers, and obeyed or exploited in communication. We see this difference as fundamental. Recanati, Adler, and Morgan & Green have their doubts. Their comments raise questions about both description and explanation. Is a speaker engaging in ostensive–inferential communication invariably presumed to obey the maxims or at least the cooperative principle? Does she invariably communicate a presumption of optimal relevance? And to what extent are

these two descriptions of the speaker equivalent? Behind these descriptive questions lies an explanatory question: Why should this be so?

Let's take the explanatory question first. According to Grice, presumptions of truthfulness, informativeness, relevance, and clarity follow from the fact that speakers are deemed to observe the cooperative principle and maxims. Where do the cooperative principle and maxims come from, and how do speakers and hearers come to know them? Grice (1975, pp. 47–49) says he does not know. By contrast, the presumption of optimal relevance is not backed by any general principle or maxim to the effect that speakers should aim at optimal relevance. It is backed by a theory of cognition from which it follows that by the very act of requesting the hearer's attention, the speaker communicates (truly or falsely) that the utterance is relevant enough to be worth the hearer's attention. So even if the two approaches were descriptively equivalent, they would differ on the explanatory level: The presumption of relevance is rooted, not in an unexplained maxim of communication, but in general and independently motivated hypotheses about cognition.

But the two approaches are not descriptively equivalent. In our approach, a particular presumption of relevance is communicated by every act of ostensive communication, whatever the communicator does; whether true or false, believed or disbelieved by the audience, this presumption plays an essential role in comprehension. On Grice's approach, the presumptions warranted by the cooperative principle and maxims are not communicated but, presumably, inferred from mutual knowledge; if the communicator blatantly violates a maxim, which she may very well do, then the corresponding presumption will not be entertained, nor will it play any role in comprehension. On the other hand, the presumption that the cooperative principle has been obeyed must be accepted as true for comprehension to work along Gricean lines.

Recanati, suggesting a parallel with the principle of relevance, says that the communicator "presents himself as conforming to the cooperative principle" and that "this is not very far from the notion that every act of communication communicates that the cooperative principle is being respected"; but his reason for saying so (as, generally, for his subtle but very idiosyncratic interpretation of the pragmatic literature) is that otherwise there would be something wrong with the Gricean account. Well, maybe there is something wrong with it. Incidentally, we would agree with Recanati that Grice's account and ours are "not very far" apart. The question is not how much they differ, but how they differ.

**3.5. Justifying the principle of relevance.** Davies asks: "From the general claim about the natural order it is said to follow that drawing someone's attention to a phenomenon 'implies a guarantee of relevance.' How exactly does this argument go?"

Four empirical claims are involved in the argument:

(i) An ostensive communicator intends the audience to pay to the ostensive stimulus the attention needed to comprehend it.

(ii) When they have the choice, people direct their attention to what is, to them, manifestly relevant enough to be worth their attention.

(iii) A particularized version of claim (i) – that is, a

version referring to a particular communicator, stimulus and audience – is mutually manifest to communicator and (attentive) audience.

(iv) Claim (ii) is a truism, mutually manifest to all normally developed humans.

The general tenor of these claims, if not their formulation, is trivial. But now remember two definitional points in our framework:

(v) An assumption is manifest if it is inferable from other manifest assumptions.

(vi) An assumption *p* is communicated if and only if it is mutually manifest to communicator and audience that the communicator intends *p* to be manifest to the audience.

It follows from (i)–(v) that it is mutually manifest to communicator and audience that the communicator intends it to be manifest to the audience that the ostensive stimulus is relevant enough to be worth the audience's attention. From this and (vi), it follows that the communicator *communicates* that the stimulus is relevant enough to be worth the audience's attention. This argument is presented, possibly in too terse a form, on page 156.

Millikan wonders whether an utterance always has to be relevant in order to be worth the hearer's attention: Could it not equally well be pleasant, amusing, or intrinsically interesting to the hearer? And is it not the case that speakers will occasionally say something for their own sake only, for example, "Please, pass the salt"?

Let us take these cases in order of increasing difficulty. First, if some information matters to the point of making you act upon it, it is surely relevant enough to be worth your attention. Requests likely to be granted are relevant in just this way. Of course, it may be more important to me to have the salt than it is to you to help me; but knowing that you are in a position to help me, and, moreover, knowing that this is mutually manifest to both of us as a result of my asking, is quite relevant to you. Being told by the one you love that you are loved (Millikan's example of intrinsic interest), even if it is the tenth time in the day, may be relevant enough because it updates and thus strengthens a very wide and accessible range of assumptions: Depending on your degree of self-confidence, the effect may be greater or lesser, but the effort of comprehension is so little!

Now consider the more difficult case of something amusing. The problem, to begin with, is that we do not know what makes, say, a joke amusing. We would like to suggest that it has to do with effect and effort, more specifically, with the fact that effort is apparently requested in vain up to the punchline, when it is rewarded by rich and unexpected effects. We do not have a developed theory to offer, but suggest that relevance theory might be extendable to aesthetics; we accept the risk of being wrong in maintaining that a proper analysis of the cases Millikan considers will show them to be no exceptions to the principle of relevance (see also para. 4.4 below).

Morgan & Green find ours "a very Panglossian view in that it treats the fact that a speaker has just uttered something communicatively as presupposing that it is the most relevant utterance that he could have uttered." Not so. We point out (on p. 157 and again on p. 162) that "The communicator may want to keep to herself the most relevant information at her disposal; she may have rea-

sons of her own for communicating information that is less relevant." The presumption communicated by every act of communication is, on the effect side, just one of adequate relevance (that is, enough to deserve attention), and not one of maximal relevance.

Morgan & Green see an empirical difference between our context-based account of relevance and one in terms of purposes and goals. In the middle of a syntax lecture, the lecturer says (truthfully), "I am allergic to chocolate." According to Morgan and Green, this utterance would count as relevant according to our theory, but be correctly predicted as irrelevant in a purpose-based account, since it "achieves no advance toward any current goal of the interchange."

In the middle of a syntax lecture, the lecturer says (truthfully), "The room is on fire." Would Morgan & Green be willing to have this remark ruled out as irrelevant on the ground that it achieves no advance toward any current goal of the interchange? We all have hundreds of longstanding goals (e.g., staying alive, remaining comfortable), which can be called up by any utterance; and new utterances can introduce new purposes and goals. The idea that any interchange has a small, preestablished set of goals is as implausible as the idea (see chap. 3, sect. 4 and 5) that the context for comprehension is fixed in advance. Given that goals can contribute to comprehension, some method of selection is needed; relevance theory provides the only one we know of that is suitable for use in anything other than the highly artificial situations considered by much work in artificial intelligence.

In the circumstances described, relevance theory would predict (correctly) that the utterance "I am allergic to chocolate" could be relevant in having contextual effects in the context of the audience's knowledge of the lecturer. Remember, though, that to be judged appropriate an utterance must be not relevant but consistent with the principle of relevance. In ordinary circumstances, "I am allergic to chocolate" would be inconsistent with the principle of relevance, unlike "The room is on fire."

Morgan & Green, who have a knack for beautiful examples, discuss cases in which silence is an ostensive stimulus. Their question is: How could silence be the most relevant ostensive stimulus the communicator could have used? Where silence conveys refusal, for example, would it not have caused the hearer less processing effort if the speaker had simply said "No"?

Morgan & Green answer their own question. To achieve an interpretation consistent with the principle of relevance, the addressee must find some contextual effects that would not be conveyed by the direct answer – effects, for example, concerning the communicator's attitude to the addressee, doubt about what she thinks or whether she should say it, and so on. As we show (pp. 196–197), any element of indirectness in an answer encourages the hearer to look for additional contextual effects to offset the extra processing effort incurred.

Bach & Harnish wonder: "How can a speaker communicate one thing by way of communicating another if the addressee stops inferring after the first thing occurs to him? More generally, how do S&W explain how the addressee can modify, supplement, or replace the first assumption that comes to mind, even if it is consistent with the PR?" First, it should be clear that the addressee does not stop at the first interpretation which occurs to



him, but at the first one which is consistent with the principle of relevance, and there indeed he stops interpreting; he does not “modify, supplement, or replace” any more.

To return to the **Morgan & Green** example of a silent refusal, why not stop when a sensible and relevant interpretation – namely: *the communicator refuses* – has been reached? Well, because that interpretation is not consistent with the principle of relevance: Some effort is left unjustified. Of course, the addressee may already have all the information he cares about. He may feel, as **Morgan & Green** suggest, that whatever else was intended is not worth the effort, and stop considering interpretations. It should be manifest to him, however, that what he has is an incomplete interpretation. Or consider an indirect speech act such as Mary’s coming back with her shopping and saying to Peter: “Oh, I forgot to buy butter and I’m too exhausted to go out again!” implicating a request that he should go and buy some butter. Why not stop at the proposition expressed? Well, again, there are circumstances in which that interpretation would not be consistent with the principle of relevance, and hence in which Mary could not have expected her utterance, on that interpretation, to be relevant enough to Peter.

And what if the communicator intended an interpretation richer than the first interpretation consistent with the principle of relevance? Too bad. She should have eliminated that first interpretation one way or another; as things stand she will be misunderstood or, at best, only partially understood. One consequence of our approach is worth noting here. As we show (p. 169) in discussing another example of silent communication, when there is a series of richer and richer conceivable interpretations of a given stimulus, the right one is the poorest interpretation consistent with the principle of relevance. It is the only interpretation the communicator can rationally *intend* to convey by means of this stimulus, however much more she may *want* to convey. This is part of what is implied by saying that human communication is inferential: The addressee might, in many cases, compute an interpretation more relevant than the first interpretation consistent with the principle of relevance by considering further implications to have been implicated by the communicator. However, he would have no rational ground for doing so. This provides a substantial constraint on interpretation, and one we have not seen mentioned elsewhere.

You know the child’s riddle: Why is it that when you have mislaid your keys it is always in the last place you look that you find them? **Adler** suggests that our argument that the first interpretation found consistent with the principle of relevance is the only one shares this riddle’s false profundity. “The first hypothesis,” he writes, “is just that hypothesis, whatever it happens to be, beyond which we are not rationally justified in going. For if we were so rationally justified, it would no longer be the first appropriate hypothesis.”

In part this is false, in part it begs the question. Suppose you are reading **Kierkegaard** and find after a while that it takes you a lot of effort to arrive at interpretations you are quite unsure of. Then it is rational for you to stop, but not to accept your interpretations as correct. Now take the case where it is rational for you to accept the interpretation you arrived at and therefore to stop search-

ing; the question remains: What made your interpretation rationally acceptable to begin with? Or to put it in other terms, suppose you are searching for your key in a drawer where there are no other keys: When you have found one, you have found yours. On the other hand, if the drawer is full of keys similar to yours, you have a problem that truisms, however disguised, won’t help you solve. Our argument shows that if you find an interpretation that is consistent with the principle of relevance and therefore *might be* the right one, then either it *is* the right one or there isn’t any right one, and that is no truism.

#### 4. Verbal communication

Several commentaries (**Blakemore**, **Carston**, **Haegeman**, **Kempson**, **Reboul**) are in the form of contributions to the development of relevance theory in the domain of verbal communication, rather than questions or criticisms, and therefore do not call for a detailed reply on our part. We greatly welcome these contributions, which is not to say that we necessarily agree with them. One aspect of the development of pragmatics which these contributions illustrate is that more and more questions which were seen as linguistic, or more generally as semiotic, and therefore to be analyzed in terms of the rules of a code, can be treated more parsimoniously in inferential terms. This raises the issue of the exact relationship between linguistics and pragmatics.

**4.1. Relevance theory and linguistics.** The obvious domain of interaction between pragmatics and linguistics is semantics, but syntax and phonology may also be concerned. **Blakemore’s** arguments for studying nontruth-conditional semantics as a distinct domain of special linguistic properties attuned to the pursuit of relevance are both important and convincing. **Haegeman** gives a very specific and clear example of a linguistic regularity which, in the framework of relevance theory, is better handled at a pragmatic than at a syntactic level. **Kempson** sees an impressive array of implications of relevance theory for the theory of grammar. The questions she raises are clearly fundamental. We do not feel competent enough, however, to evaluate her answers until they are further developed.

At the phonological level, there is a well-known interaction between stress patterns and pragmatic effects (see chap. 4, sect. 5). On the other hand, it might seem that standard segmental phonology is not affected by pragmatic considerations, and that is what **Cutler** takes us to imply, not just about phonology but about the whole linguistic decoding level when we stress its automatic character. She then shows beautifully how speakers encode their messages as to spare their hearers unnecessary decoding effort, however minute.

To say that a process is automatic is not, however, to “dismiss” it, as **Cutler** says we do with linguistic decoding, nor is it to say that the amount of effort the process requires is beyond the speaker’s control: It can be controlled by controlling the input. The level of effort which affects the relevance of an utterance is not just that of the inferential process which takes as input the output of the decoding process: It is that of the whole processing of the utterance, including decoding. **Cutler** writes that

“speakers’ attention to ensuring relevance is merely one end of a continuum of hearer-coddling; there is certainly nothing special about it, and nothing that makes attention to hearers’ inference processes qualitatively different from attention to hearers’ decoding processes.” From our point of view, speakers’ attention to hearers’ decoding processes is part of their attention to ensuring relevance; it contributes to relevance by diminishing effort and in that sense is not qualitatively different from attention to hearers’ inference processes. The only point at which we disagree with Cutler is when she writes as if we would disagree with her.

**4.2. Relevance, explicature, and implicature.** In *Relevance* we define explicitness in a nonconventional way: “An assumption communicated by an utterance *U* is *explicit* if and only if it is a development of a logical form encoded by *Y*” (p. 182). We do not define “development,” and we should have done so: the omission confuses Levinson. But we do discuss and illustrate at length what we have in mind: the inferential completion of a decoded logical form by identifying referents, by semantic enrichment, and by specifying propositional attitudes. We have coined the term “explicature” to refer to an explicitly communicated assumption, and we redefine “implicature” as a nonexplicitly communicated assumption. Levinson complains that the explicature/ implicature distinction “is not clear. The only criterion offered is that explicatures must contain the encoded SR [semantic representation] or LF [logical form] as a proper subpart.” He then gives an example of what we would agree is an implicature that nevertheless meets this criterion.

To answer Levinson, let us take an even simpler example. Suppose that (6b), which explicitly communicates (7), implicitly communicates (8):

- (6) (a) *Mary*: Who bought this car?
- (b) *Peter*: John did.
- (7) John bought this car.
- (8) John did something silly.

Now (8) contains the logical form of (6b) and yet is implicit. Is this a counterexample to our definition? No, because our criterion is the *development* of the logical form, not merely its implicature. (8) is arrived at purely inferentially whereas (7), the true explicature of (6b), results from the inferential development of the decoded logical form of (6b). Many of the implications of the explicature/implicature distinction drawn this way are cogently presented in Carston’s commentary.

According to Levinson, an understanding of generalized conversational implicatures is the signal achievement of neo-Gricean pragmatics, but *Relevance* hardly engages this work.

Grice (1975, p. 56) distinguished between generalized implicatures, those “normally carried by saying that *p*,” and particularized implicatures, “carried by saying that *p* on a particular occasion in virtue of special features of the context.” His best-known examples are particularized implicatures; the discussion of generalized implicatures is restricted to a few cases; and there is no evidence that he saw the distinction as theoretically significant.

“Generalized implicature” and “particularized implicature” are not classificatory theoretical concepts. They do not define distinct natural kinds but lie at

different ends of a spectrum, with a fuzzy borderline area in between. The methods by which generalized and particularized implicatures are created must be the same, and the same theory must account for both. The danger of concentrating on generalized implicatures is that one may end by inventing specialized procedures which deal more or less well with central cases, but do not extend to particularized implicatures at all. The equation of “generalized implicatures” with “default inferences” seems to us to run this risk.

According to Levinson, the theory of generalized implicatures makes predictions about the following example which differ from those made by relevance theory:

- (9) A: If it’s possible that the spy has *more* than two passports, he may yet escape.
- B: He has two passports.

The theory of generalized implicatures predicts that B implicates that the spy has only two passports because the use of “two” normally implicates “at most two.” Levinson believes relevance theory would predict “that there would be no explicature ‘at most two’, because that would rule out the low-cost contextual implication ‘he may escape,’ thus lowering R[ellevance]. Thus B’s utterance should implicate ‘He may escape.’” Since this is counterintuitive, he concludes that his theory is better than ours.

This is a total misapplication of relevance theory, which does not predict that the most relevant interpretation conceivable is the right one. Moreover, “the spy has (at least) two passports” does not imply “it is possible that he has more than two” so that, with or without the “at most” implicature, B’s utterance does not normally implicate that the spy may escape. It would carry this implicature only in the odd case in which the assumption “If the spy has two passports, then it is possible that he has more than two” were manifest. Otherwise, two types of cases must be considered: If it is manifest both that B knows exactly how many passports the spy has and that he is willing to share the information with A, then B’s utterance will communicate (by enrichment of the explicature rather than by implicature) that the spy has exactly two passports. If either of these two conditions is not met, B’s utterance would indicate, and might implicate, that B is unable or unwilling to say whether the spy has more than two passports. Thus, relevance theory predicts a range of possible interpretations, the plausibility of which depends on that of the situations in which they occur.

For Levinson, the best approach to this example is to assign it a “default” interpretation which must somehow be overridden in cases in which it is not in fact used. We see no reason to follow him. For both generalized and particularized implicatures, the correct interpretation is determined by an interaction between contextual assumptions, linguistic form, and some pragmatic criterion. A theory which sets out to concentrate only on particular aspects of linguistic form, or particular types of context, is in danger of being trapped by the stereotypes it should be studying and might miss significant generalizations over the whole domain.

Where do interpretive hypotheses come from? ask **Bach & Harnish**. C. Smith makes the question much more precise by analyzing a few linguistic examples and showing in detail the different sources of the information

needed in our account of inferential comprehension. She concludes, however, that our account is not detailed enough to “work out how information is accessed and how assumptions are constructed; in short, to see how the processes of inference work.” Where do implicated premises come from? asks Levinson, even more specifically. They cannot be deduced; nor can they always be retrieved, since an implicated premise may contain information that is new to the hearer.

At the general level envisaged by Bach & Harnish, it is true that our speculative remarks (pp. 82–83, and chap. 4, sect. 2) do not provide a satisfactory answer to the question of where hypotheses come from. However, this question is not an artifact of relevance theory but a very general problem of cognitive psychology. As we point out, “Since relevance theory is, among other things, an attempt to ground models of human communication squarely in cognitive psychology, it cannot just take advantage of the insights of cognitive psychology, but must also share its weaknesses” (p. 170). It seems to us, however, that, when the issues are considered in detail, relevance theory does illuminate them. To each of the specific questions raised by C. Smith there are, in *Relevance*, at least fragments of an answer, which she does not discuss. We will therefore restrict our discussion to the case of implicated premises raised by Levinson.

Let us, to begin with, distinguish retrievability and accessibility. Information is retrievable if it is explicitly represented in memory. Information that is not retrievable may nonetheless be accessible: from the environment, by perception; from memory, by deduction, non-demonstrative inference, or the development of assumption schemas; and by the type of heuristic described (pp. 82–83), by which, on presentation of assumptions of a certain form, one considers whether one has evidence for assumptions of a related form. For example, given an assumption of the form *If P then Q*, a reasonable heuristic would be to consider whether one has evidence for an assumption of the form *P*; given an assumption of the form *All F are G*, a reasonable heuristic (and hence one that a speaker might expect a hearer to use) would be to consider whether one has evidence for an assumption of the form *x is F*.

Consider the exchange discussed by Levinson:

- (10) (a) *Peter*: Would you drive a Zordia?  
 (b) *Mary*: I wouldn't drive ANY expensive car.

On hearing Mary's reply to his question, a reasonable heuristic for Peter would be to consider whether he has any evidence for the assumption in (11):

- (11) A Zordia is an expensive car.

Let us suppose that before Mary spoke he had no evidence for it. We claim (pp. 116, 195) that Mary's utterance will provide it. Her utterance, let us further suppose, is consistent with the principle of relevance only if (11) is added to the context as an assumption of a certain degree of strength. It follows that Mary must have expected Peter to use that assumption at least at that degree of strength. That is, Mary's utterance provides Peter with indirect evidence that she believes (11); and if he trusts her enough, he may come to believe it too. In Wilson and Sperber (1986a), many of Grice's examples of implicature, and other standard cases, for example bridging

implicature, are reanalyzed along these lines (though in an earlier framework than the one developed in *Relevance*).

**4.3. Relevance, tropes, and interpretive use.** An important difference between Gricean pragmatics and relevance theory is that a presumption of optimal relevance does not entail a presumption of literal truthfulness. Noting this, Levinson, Adler and McCawley suggest that our framework should be supplemented with a maxim of truthfulness. Yet in *Relevance* (chap. 4, sect. 8) and in Sperber and Wilson (1986b) we argue that the maxim of truthfulness is too strong. To adopt such a maxim is to commit oneself to analysing metaphor and irony, loose talk, and free indirect speech, for example, as deviations from the norm, or at least from the favoured interpretation. And accounts along these lines, we argue, are both descriptively and explanatorily inadequate.

The account offered in *Relevance* (chap. 4, sect. 7–9) is superior from both descriptive and explanatory points of view. We show that the presumption of relevance entails something much weaker than a presumption of literal truthfulness, and weaker in just the right way. Utterances purport to be optimally relevant representations of the speaker's thoughts – hopes, fears, desires, beliefs, fantasies, and so on. An utterance can be optimally relevant without being literally truthful: For example, it may represent not a belief of the speaker's but a belief attributed to someone else (as in the case of irony and free indirect speech); and it may merely resemble, rather than literally reproduce, the belief it represents (as in the case of metaphor and loose talk). Seen from this perspective, literal truthfulness is just one way of achieving optimal relevance, adequate on some occasions only. It needs no special maxim to explain it, and its existence does not make metaphor and irony any harder to explain.

Our accounts of metaphor and irony are based on a notion of representation by resemblance, or interpretive use, which also sheds light on the problem of the expression of propositional attitudes. Recanati believes that an adequate account of the same phenomena can be developed without recourse to the notion of interpretive use. The *communicatum*, he says, may bear one of several relations  $R^*$  to the proposition  $p$  expressed by an utterance: It may be identical to  $p$  (as in ordinary assertions); it may be an embedding of  $p$  under some higher predicate (as in irony and free indirect speech); or it may be a subset of the “entailments” of  $p$  (as in metaphor and loose talk).

The problem with Recanati's account is that, even if it could be made descriptively adequate, unless it were backed by the notion of interpretive use it would still not be explanatory. What determines the class of relations  $R^*$  which the *communicatum* may bear to the proposition expressed? Why just this list of relations? Why could the *communicatum* not be a conjunction or disjunction of  $p$  with some contextually salient proposition, or a conditional with  $p$  as antecedent and some contextually salient proposition as consequent? How does the child acquire knowledge of  $R^*$ ? What is needed is an explanatory account of  $R^*$  itself.

In our framework, such an account is given by the notion of interpretive use. Given that any object in the world may be used to represent any other object it resembles, it is no surprise to find that utterances may be used for their interpretive resemblance to thoughts of the



speaker, which may themselves be interpretive representations of thoughts attributed to someone else. The difference between literal and nonliteral assertions is simply a difference in degree of resemblance between the proposition expressed by the utterance and the thought of the speaker; the difference between literal and ironical utterances is simply a difference between attributive and nonattributive utterances. The notion of interpretive use explains phenomena which **Recanati's** account at best describes.

**Gibbs** believes that in our model a nonliteral interpretation would take more processing effort than a literal one. (He points out that this is contradicted by much recent research.) **Recanati** seems to believe something similar, but he draws an opposite conclusion – namely, that we should assume that a literal interpretation is preferred, or, as he puts it, that there is a “presumption of literalness.”

Both **Gibbs** and **Recanati** make an erroneous assumption: Our model does not predict that a literal interpretation requires less processing effort; if anything, it predicts the opposite. A literal interpretation (in a case of “saying that”) is one in which the speaker guarantees the truth of the proposition literally expressed. In the comprehension process, as we describe it, the hearer starts computing contextual implications even before the sentence has been entirely uttered, stopping when the set of implications so computed allows him to identify an interpretation consistent with the principle of relevance. If accepting that interpretation as true makes manifest all the implications, analytic or contextual, of the proposition expressed, then the interpretation is literal; otherwise it is not. In most cases, the interpretation arrived at does not carry all the implications of the proposition expressed, but only a subset of them. Those implications of the proposition expressed that are not implied by the interpretation arrived at need not have been rejected: They may not have been considered at all.

If we followed **Recanati's** advice and introduced a presumption of literalness, then enough implications would have to be considered for the literal interpretation to be accepted or rejected. The amount of effort required for comprehension would be systematically greater than we claim; nonliteral interpretations would require even more processing effort; and **Gibbs** would be right in using against our account evidence which we see, on the contrary, as very much in its favour.

**Adler** suggests, as an alternative to our proposal, that **Grice's** account of irony should be supplemented with a condition requiring mutual knowledge that there exists a practice of ironical usage (i.e., of saying one thing and implicating the opposite).

**Grice** himself (1978, p. 124) is dissatisfied with his own account of irony as saying one thing and implicating the opposite (a variant of the classical view of irony as saying one thing and figuratively meaning the opposite).<sup>5</sup> What is missing from both accounts, we argue, is that irony is echoic: The speaker echoes and dissociates herself from an opinion implicitly attributed to someone else. This proposal is empirically distinguishable from the classical view in that a speaker can ironically echo an opinion she finds absurd or irrelevant without necessarily indicating that she believes the opposite, a possibility illustrated by many ironical quotations, exclamations and understatement.

Moreover, **Adler's** proposed view of tropes as a matter of convention or social practice – (a view, incidentally, that is rejected by **Grice** (1978, p. 124) – is not explanatory: Why the same tropes in culture after culture? Why not a trope based on permutation of subject and object, or of the referents of “I” and “you”? The account in *Relevance* suggests an explanation: Metaphor and irony are special cases of representation by resemblance, which, as we show, is universally exploited in both verbal and nonverbal communication.

**N. V. Smith** has asked us to draw attention to any misapprehensions in his comments on interpretive use, on the ground that they are (a) likely to be shared by others and (b) our fault. Well, first, there is not, as he suggests, an equivocation in our uses of “interpretation” in Figure 3 (p. 232). In both cases, the term refers to resemblances between propositional forms: of an utterance and a thought of the speaker's on the one hand, and of a thought of the speaker's and an attributed or desirable thought on the other.

In a sense, all utterances could be said to be interpretive in that they represent thoughts of the speaker. However, we talk of an utterance as descriptive when the speaker's thought is itself descriptive, and as interpretive only when the speaker's thought is itself interpretive.

Second, the existence of ironical and, more generally, echoic imperatives and interrogatives is not incompatible with figure 3. The fact that attributed thoughts may themselves be descriptive or interpretive introduces an element of recursion (as mentioned on p. 231). Consider (12):

(12) Don't mind me.

This utterance has at least two interpretations. Mary might be expressing her own belief that it's all right for Peter to ignore her. This interpretation requires just one pass through the diagram, taking path (d). Or Mary might be speaking ironically, attributing to Peter the thought that it's all right for him to ignore her, and dissociating herself from that thought. This interpretation requires (at least) two passes through the diagram: the first, taking path (a), to indicate that the utterance is a second-degree interpretation; the second, taking path (d), to indicate the properties of the attributed thought. If the attributed thought is itself attributive (e.g., if Mary attributes to Peter the thought that *she* thinks it is all right for him to ignore her), a still further pass is needed.

Third, ironical utterances and utterances marked with “hearsay” particles are interpretive in just the same sense: Both are second-degree interpretations. What distinguishes irony from regular “hearsay” utterances is not the type and degree of interpretiveness involved but the presence or absence of a certain type of attitude. In irony, the speaker dissociates herself from the opinion echoed. As **Blass** (1987) shows, utterances marked with hearsay particles need not be accompanied by such dissociative attitudes; but they can be, and the result is irony. The use of hearsay particles for irony – entirely puzzling on the classical account – is striking confirmation of our analysis of irony.

Fourth, as the text was supposed to make clear, example (104) “If giraffes had wings, they would lay eggs” (p. 229) is a case of interpretive as opposed to descriptive use

not because it is a conditional but because it is an example, used to represent an utterance type or assumption not attributed to anyone, but worth considering for its intrinsic properties.

**4.4. Relevance and fiction.** Clark believes that relevance theory does not apply to fiction because it cannot handle “layering,” the phenomenon whereby communication takes place simultaneously at different levels. Reboul believes that, on the contrary, relevance theory, and more particularly its treatment of metaphor, offers an adequate (and novel) approach to fiction. We agree with Reboul that relevance theory is relevant to the study of fiction, and with Clark that layering is essential to the issue. We believe moreover that relevance theory is uniquely well equipped to handle layering. As we define it, a communicative intention is an informative intention about an informative intention, making it possible for a communicative intention itself to be the object of a second-order (or an *n*th-order) communicative intention. In other words, a first-level act of ostensive communication can serve as an ostensive stimulus for a second-level act of ostensive communication.

A good example of layering is provided by Morgan & Green: Mozart says to Salieri, “I never thought music like that was possible” in a context such that Salieri is incapable of resolving the equivocation and deciding whether to take the remark as an insult or as a compliment. Morgan & Green believe this case should pose problems for us because they take the principle of relevance to be incompatible with deliberate ambiguity. However, deliberate ambiguity at one level can be used as a nonambiguous ostensive stimulus at another level. By putting Salieri in a situation in which he cannot tell whether he is being complimented or insulted, Mozart makes it manifest that there is much less mutual understanding between them than Salieri might wish. Moreover, Mozart does so in a manifestly intentional way: The failure of the first-level communication successfully communicates Mozart’s sense of distance from Salieri on a second level.

There are cases when successful first-level communication is complemented by equally successful second-level communication. Good fiction is a case in point. As you read, say, a novel, each sentence, whether descriptive of situations or interpretive of the characters’ thoughts and words, is relevant in a context essentially provided by preceding sentences and general knowledge; it is understood on the basis of the principle of relevance in the usual way. But there is a discontinuity between the world described in a work of fiction and the world of the reader such that the relevance of the work as a whole is not immediately obvious. As Reboul puts it: Why are people prepared to spend time on fiction, when they know it is false? We are not convinced that extending the idea of a metaphor, understood as an interpretation of the author’s thoughts, really answers the question. We believe that the notion of second-order communication (Clark’s layering), together with the notions of weak implicature and of *showing* as a form of communication, may be more helpful.

Contextual effects come not only in the form of new contextual implications, but also in the form of the strengthening of previously manifest assumptions. Relevance can be achieved not only by informing someone of

new facts, but also by altering saliences and strengths in the cognitive environment. This is, we suggest, the way in which many examples are relevant: They force the listener or reader to develop or otherwise modify mental models, scenarios, scripts, or schemas. For examples to be relevant in this way, it does not matter whether they are true or fictional provided that they are recognizably, in some respects at least, truthlike. So communicating, at one level, information about Hamlet or Ishmael which is relevant only in its own artificial context communicates at a second level by *showing* what is possible or conceivable, rather than what is.

Clark is right: What Shakespeare and Melville communicate to us through their work is not a Gricean speaker’s meaning. But Clark is also wrong: These works are perfectly good cases of ostensive communication because they are, at the higher level, cases of showing rather than of “saying that.” They achieve relevance through a vast array of weak effects, rather than through a “meaning” or a “message.”

#### NOTES

1. Here Bach & Harnish, referring to the speaker’s informative intention, oddly add parenthetically: (“if he even has one”). By our definition, a communicator necessarily has an informative intention, since communication consists in making such an *informative intention mutually manifest*. What a communicator may lack is a *communicative* intention (see pp. 63–64).

2. A further possible constraint is mentioned in Chapter 2, footnote 23 and discussed in Chapter 4, Section 5.

3. Contrary to the claim Hinkelman attributes to us, the deductive device does not compute “the most relevant implicatures [i.e., implications] first.”

4. Wilks misrepresents our view to such an extent that correction rather than discussion is called for. He describes our work as remaining “within the process-free linguistic paradigm” while in fact what we outline is, good or bad, a process model, as emphasized by other commentators. He imagines that we “do not really accept that such inferences [as occur in utterance interpretation] must be some particular individual’s inferences, and so [we] feel free to opt for an ‘objective’ set of hypothesized premises, ones [we] believe to be correct.” In reality, we are the first in the literature to develop a concept of relevance *to an individual* (chap. 3, sect. 5), which is essential to our whole theory.

5. Grice gives an example of his own designed to show that all his conditions for irony could be met without irony being perceived. Our example (113) (on p. 240), discussed by Adler, was designed to make a different though related point: The failure to perceive irony cannot be put down to a lack of mutual knowledge or shared assumptions, because even when the lack is repaired and the intended interpretation spelled out, the utterance is still not acceptable as irony.

#### References

- Allen, J. (1983) Recognizing intentions from natural language utterances. In: *Computational models of discourse*, ed. M. Brady & R. Berwick. MIT Press. [SCL]
- (1987) *Natural language processing*. Benjamin Cummings. [EH]
- Armstrong, D. (1971) Meaning and communication. *Philosophical Review* 80:427–47. [aDS]
- Atlas, J. & Levinson, S. (1981) *It-clefts, informativeness and logical form*. In: *Radical pragmatics*, ed. P. Cole. Academic Press. [SCL]
- Austin, J. L. (1962) *How to do things with words*. Oxford University Press. [HHC]

- Bach, K. & Harnish, R. M. (1979) *Linguistic communication and speech acts*. MIT Press. [aDS, KB]
- Bagley, W. A. (1900) The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology* 12:80–130. [AC]
- Bennett, J. (1976) *Linguistic behaviour*. Cambridge University Press. [aDS]
- Blackburn, S. (1984) *Spreading the word*. Oxford University Press. [aDS, MD]
- Blakemore, D. L. (1987) *Semantic constraints on relevance*. Blackwell. [DB]
- Blass, R. (1987) The grammaticalisation of interpretive use: the case of *re* in Sissala. Paper presented to the spring meeting of the Linguistics Association of Great Britain. [NVS, rDS]
- Bochenski, I. M. (1947) *La Logique de Théophraste* (Collectanea Friburgensia, ms. fasc. 37). Fribourg. [YW]
- Bond, Z. S. & Small, L. H. (1983) Voicing, vowel and stress mispronunciations in continuous speech. *Perception & Psychophysics* 34:470–74. [AC]
- Bransford, J. & McCarrell, N. (1977) A sketch of a cognitive approach to comprehension. In: *Thinking*, P. Johnson-Laird & P. Wason, ed. Cambridge University Press. [rDS]
- Bretherton, I., McNew, S. & Beeghly-Smith, M. (1981) Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind"? In: *Infant social cognition*, ed. M. E. Lamb & L. R. Sherod. Erlbaum. [rDS]
- Brown, G. & Yule, G. (1983) *Discourse analysis*. Cambridge University Press. [aDS]
- Bruce, B. (1981) A social interaction model of reading. *Discourse Processes* 4:273–311. [rDS, HHC]
- Carnap, R. (1950) *Logical foundations of probability*. Routledge & Kegan Paul. [aDS]
- Carston, R. (forthcoming) Explicature, implicature and truth-theoretic semantics. In: *Mental representations: The interface between language and reality*, ed. R. Kempson. Cambridge University Press. [RC]
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and points of view. In: *Subject and topic*, ed. C. Li. Academic Press. [aDS]
- Charniak, E. (1986) A neat theory of marker passing. In *Proceedings of the American Association of Artificial Intelligence, 1986*. Morgan Kaufmann. [EH]
- Chomsky, N. (1981) *Lectures on government and binding*. Foris. [RK]
- (1986) *Knowledge of language: Its nature, origin and use*. Praeger. [RK]
- Clark, H. H. (1977) Bridging. In: *Thinking: Readings in cognitive science*, ed. P. Johnson-Laird & P. Wason. Cambridge University Press. [aDS]
- (1977a) Inferences in comprehension. In: *Basic processes in reading: Perception and comprehension*, ed. D. Laberge & S. Samuels. Erlbaum. [RWG]
- (1982) The relevance of common ground: Comments on Sperber and Wilson. In: *Mutual knowledge*, ed. N. V. Smith. Academic Press. [HHC]
- (1987) Four dimensions of language use. In: *The pragmatic perspective: Proceedings of the 1985 International Pragmatics Conference*, ed. M. B. Papi & J. Verschueren. John Benjamins. [rDS, HHC]
- Clark, H. H. & Carlson, T. (1981) Context for comprehension. In: *Attention and performance IX*, ed. J. Long & A. Baddeley. Erlbaum. [aDS]
- Clark, H. H. & Gerrig, R. J. (1984) On the pretense theory of irony. *Journal of Experimental Psychology: General* 113:121–26. [HHC]
- Clark, H. H. & Haviland, S. (1977) Comprehension and the given-new contract. In: *Discourse production and comprehension*, ed. R. Freedle. Ablex. [aDS]
- Clark, H. H. & Marshall, C. R. (1981) Definite reference and mutual knowledge. In: *Elements of discourse understanding*, ed. A. K. Joshi, B. L. Webber & I. A. Sag. Cambridge University Press. [rDS, RJG]
- Cole, R. A. (1973) Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics* 11:153–56. [AC]
- Cooper, W. E. & Paccia-Cooper, J. (1980) *Syntax and speech*. Harvard University Press. [AC]
- Cottrell, G. W. (1985) *A connectionist approach to word sense disambiguation*. University of Rochester Department of Computer Science, TR 154. [EH]
- Cummins, R. (1983) *The nature of psychological explanation*. MIT Press. [rDS, PP]
- Cutler, A. (1980) *Productivity in word formation. Papers from the Sixteenth Regional Meeting, Chicago Linguistic Society*. [AC]
- (1983) Speakers' conceptions of the functions of prosody. In: *Prosody: Models and measurements*, ed. A. Cutler & D. R. Ladd. Springer. [AC]
- (1986) *Forbear* is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29:201–20. [AC]
- Dascal, M. (1977) Conversational relevance. *Journal of Pragmatics* 1:309–28. [SCL]
- Davidson, D. (1978) What metaphors mean. *Critical Inquiry* 5:31–47. [MD]
- (1984) Communication and convention. In: *Truth and interpretation*, by D. Davidson. Clarendon Press. [aDS]
- Davies, M. (1981) *Meaning, quantification, necessity: Themes in philosophical logic*. Routledge & Kegan Paul. [aDS]
- (1983) Idiom and metaphor. *Proceedings of the Aristotelian Society* 83:67–85. [MD]
- de Beaugrande, R. & Dressler, W. (1981) *Introduction to text linguistics*. Longman. [aDS]
- Dell, G. S. (1986) A spreading activation theory of retrieval in sentence production." *Psychology Review* 93: 283–321. [EH]
- Derbyshire, D. (1985) *Hixkaryana and linguistic typology*. Publications in Linguistics No. 76, Summer Institute of Linguistics, Texas. [NUS]
- Donnellan, K. (1966) Reference and definite descriptions. *Philosophical Review* 75:281–304. [JDM]
- Dowty, D., Wall, R. & Peters, S. (1981) *Introduction to Montague semantics*. Reidel, Dordrecht. [RK]
- Evans, G. & McDowell, J. (1975) *Introduction to Truth and meaning: Essays in semantics*. Oxford University Press. [PP]
- Fodor, J. (1983) *The modularity of mind*. MIT Press. [aDS, RK]
- (1985) Précis of *The modularity of mind*. *Behavioral and Brain Sciences* 8:1–42. [aDS]
- Frisch, A. M. (1985) Using model theory to specify AI programs. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, ed. A. Joshi. Morgan Kaufmann. [EH]
- Gazdar, G. (1979) *Pragmatics: Implicature, presupposition and logical form*. Academic Press. [aDS, RC, SCL]
- Gazdar, G. & Good, D. (1982) On a notion of relevance. In *Mutual knowledge*, ed. N. Smith. Academic Press. [SCL]
- Gibbs, R. (1984) Literal meaning and psychological theory. *Cognitive Science* 9:275–304. [RWG]
- (1986) On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General* 115:3–15. [RWG]
- (1987) *Mutual knowledge and the psychology of conversational inference*. *Journal of Pragmatics* 11. [RJG, RWG]
- Givón, T., ed. (1979) *Syntax and semantics 12: Discourse and syntax*. Academic Press. [aDS]
- Goffman, E. (1974) *Frame analysis*. Harper & Row. [rDS, HHC]
- Goodman, N. (1968) *Languages of art*. Oxford University Press. [PP]
- Green, G. & Morgan, J. (1981) Pragmatics, grammar and discourse. In: *Radical Pragmatics*, ed. P. Cole. Academic Press. [aDS]
- Grice, H. P. (1957) Meaning. *Philosophical Review* 66:377–88. Reprinted (1971) in *Semantics: An interdisciplinary reader*, ed. D. Steinberg & L. Jakobovits. Cambridge University Press. [aDS]
- (1968) Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language* 4:225–42. Reprinted (1971) in *The philosophy of language*, ed. J. Searle. Oxford University Press. [aDS, RWG]
- (1969) Utterer's meaning and intentions. *Philosophical Review* 78:147–77 [aDS, KB]
- (1975a) Logic and conversation. In: *Syntax and semantics 3: Speech acts*, ed. P. Cole & J. Morgan. Academic Press. [aDS, DB, HHC, RC, SCL, JLM]
- (1975b) Logic and conversation. In: *The logic of grammar*, ed. D. Davidson & G. Harman. Dickenson Publishing. [JEA]
- (1978) Further notes on logic and conversation. In: *Syntax and semantics 9: Pragmatics*, ed. P. Cole. Academic Press. [aDS, HHC, RC]
- (1982) Meaning revisited. In: *Mutual knowledge*, ed. N. Smith. Academic Press. [aDS]
- Grimshaw, J. (1979) Complement selection and the lexicon. *Linguistic Inquiry* 10(2):279–326. [aDS]
- Grosz, B. J. & Sidner, L. L. (1986) Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3): 175–204. [EH]
- Gussenhoven, C. (1983) Focus, mode and nucleus. *Journal of Linguistics* 19(2): 377–418. [RK]
- Halliday, M. (1967–8) Notes on transitivity and theme in English. *Journal of Linguistics* 3:37–81, 199–244; 4:179–215. [aDS]
- Halpern, J. & Moses, Y. (1985) A guide to the modal logics of knowledge and belief. In *Proceedings of the ninth international joint conference on artificial intelligence, Los Angeles*. Morgan Kaufmann. [rDS, SJR]
- Harman, G. (1968) Three levels of meaning. *Journal of Philosophy* 65:590–602. Reprinted (1971) in *Semantics: An interdisciplinary reader*, ed. D. Steinberg and L. Jakobovits. Cambridge University Press. [aDS]
- Hintikka, J. (1962) *Knowledge and belief*. Cornell University Press. [SJR]
- Holdcroft, D. (1987) Conversational relevance. In: *The pragmatic perspective*, ed. M. Papi & J. Verschueren. Benjamin. [SCL]
- Horn, L. (1972) *On the semantic properties of the logical operators in English*. Indiana University Linguistics Club. [SCL]
- (1984) Towards a new taxonomy for pragmatic inference: Q-based and R-based implicature. In: *Georgetown University round table on languages*



- and linguistics 1984, ed. D. Schiffrin. Georgetown University Press. [RC, SCL, JDM]
- (1985) Metalinguistic negation and pragmatic ambiguity. *Language* 61(1):121–74. [SCL]
- Jackendoff, R. (1972) *Semantic interpretation in generative grammar*. MIT Press. [aDS]
- Johnson-Laird, P. (1983) *Mental models*. Cambridge University Press. [aDS]
- Jorgensen, J., Miller, G. & Sperber, D. (1984) Test of the mention theory of irony. *Journal of Experimental Psychology: General* 113(1):112–20. [aDS]
- Kaisse, E. M. (1985) *Connected speech: The interaction of syntax and phonology*. Academic Press. [AC]
- Kanal, L. N. & Lemmer, J. F., eds. (1986) *Uncertainty in artificial intelligence*. Elsevier. [SJR]
- Karttunen, L. (1976) Discourse referents. In: *Notes from the linguistic underground*, ed. J. McCawley. Academic Press. [JDM]
- Kempson, R. (1986) Definite NPs and context-dependence: A unitary theory of anaphora. In: *Reasoning and discourse processes*, ed. T. Myers. Academic Press. [RK]
- (forthcoming a) On the interface of grammar and cognition: The principle of full interpretation. In: *Mental representations: The language reality interface*, ed. R. Kempson, Cambridge University Press. [RK]
- (forthcoming b) Reconstruction and logical form. *Linguistics*. [RK]
- Kempson, R. & Cormack, A. (1981). Ambiguity and quantification. *Linguistics & Philosophy* 4:259–309. [SCL]
- Konolige, K. (1986) *A deduction model of belief*. Pitman. [SJR]
- Leech, G. (1983) *Principles of pragmatics*. Longman. [aDS]
- Leslie, A. M. (1987) Pretense and representation: The origins of “theory of mind.” *Psychological Review* 94(4):412–46. [rDS]
- Levelt, W. J. M. & Cutler, A. (1983) Prosodic marking in speech repair. *Journal of Semantics* 2:205–17. [AC]
- Levinson, S. (1983) *Pragmatics*. Cambridge University Press. [aDS, RC, SCL, JDM]
- Lewis, D. (1969) *Convention*. Harvard University Press. [arDS, RJG]
- (1975) Languages and language. In: *Language, mind and knowledge*, ed. K. Gunderson. Minnesota University Press. Reprinted (1983) in vol. 1 of D. Lewis, *Philosophical papers*. Oxford University Press. [aDS]
- (1979) Scorekeeping in a language game. *Journal of Philosophical Logic* 8:339–59. Reprinted (1983) D. Lewis, in *Philosophical papers*. Oxford University Press. [aDS, JDM]
- Loar, B. (1976) Two theories of meaning. In: *Truth and meaning*, ed. G. Evans & J. McDowell. Oxford University Press. [aDS]
- (1981) *Mind and meaning*. Cambridge University Press. [aDS]
- Lyons, J. (1977) *Semantics*, vols. 1 & 2. Cambridge University Press. [aDS]
- Macnamara, J. (1982) *Names for things: A study of human learning*. Bradford Books MIT Press. [JM]
- (1986) *A border dispute: The place of logic in psychology*. Bradford Books/MIT Press. [JM]
- Marslen-Wilson, W. D. & Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10:29–63. [AC]
- McCawley, J. D. (1979) Presupposition and discourse structure. In: *Presupposition*, ed. D. Dinneen & C.-K. Oh. Academic Press. [JDM]
- (1981) *Everything that linguists have always wanted to know about logic (but were ashamed to ask)*. University of Chicago Press. [JDM]
- (1985) Actions and events despite Bertrand Russell. In: *Actions and events*, ed. E. LePore & B. P. McLaughlin, Basil Blackwell. [JDM]
- McDowell, J. (1980) Meaning, communication and knowledge. In: *Philosophical subjects*, ed. Z. van Straaten. Clarendon Press. [aDS, MP, PP]
- Millikan, R. G. (1984a) *Language, thought and other biological categories*. Bradford Books/MIT Press. [RCM]
- (1984b) Naturalist reflections on knowledge. *Pacific Philosophical Quarterly* 65:315–334. [RCM]
- Norvig, P. (1987) A unified theory of inference for text understanding. Ph.D. dissertation, Computer Science Division, University of California, Berkeley. [SJR]
- Oh, C.-K. & Dinneen, D., eds. (1979) *Syntax and semantics 11: Presupposition*. Academic Press. [aDS]
- Partee, B. (1984) Compositionality. In: *Varieties of formal semantics*, ed. F. Landman & F. Veltman. Foris, [RK]
- Patton, T. & Stampe, D. (1969) The rudiments of meaning: On Ziff on Grice. *Foundations of Language* 5(1):2–16. [aDS]
- Peirce, C. S. (1955) Abduction and induction. In: *Philosophical writings of Peirce*, ed. J. Buchler. Dover [YW]
- Perrault, C. R. (1987) An application of default logic to speech act theory. In: *SDF Benchmark Series: Plans and intentions in communication and discourse*, ed. P. R. Cohen, J. Morgan & M. Pollack. MIT Press. [rDS, EH]
- Perrault, R. & Allen, J. (1980) A plan-based analysis of indirect speech acts. *American Journal of Computer Linguistics* 6:167–82. [YW]
- Pesetsky, D. (1985) Morphology and logical form. *Linguistic Inquiry* 16(2) 193–246. [RK]
- Prince, E. (1978) On the function of existential presupposition in discourse. *Papers from the 14th Regional Meeting*, Chicago Linguistic Society. [JDM]
- (1981) Towards a taxonomy of given-new information. In: *Radical pragmatics*, ed. P. Cole. Academic Press. [aDS]
- Quine, W. V. O. (1960) *Word and object*. MIT Press. [SJR]
- Reboul, A. (1986) L'interprétation des énoncés de fiction. *Cahiers de Linguistique française* 7:27–41. [AB]
- Recanati, F. (1979) *La transparence et l'énonciation*. Paris: Seuil. [aDS]
- (forthcoming) *Performative utterances*. Cambridge University Press. [aDS]
- Reinhart, T. (1981) Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27:53–94. [aDS]
- Rizzi, L. (1986) Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17(3):501–557. [LH]
- Rochemont, M. (1986) *Focus in generative grammar*. Benjamin. [aDS]
- Russell, B. (1905) On denoting. *Mind* 14:479–93. [JDM]
- Sadock, J. (1985) Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language and Linguistic Theory* 3:379–439. [RK]
- (1986) Remarks on the paper by Wilson and Sperber. *Papers from the Chicago Linguistic Society, Part 2*: 85–90. [SCL]
- Sadock, J. & Zwicky, A. (1985) Speech act distinctions in syntax. In: *Language typology and syntactic description*, vol. 1, ed. T. Shopen. Cambridge University Press. [aDS]
- Sag, I. (1981) Formal semantics and extralinguistic context. In: *Radical pragmatics*, ed. P. Cole. Academic Press. [aDS]
- Schank, R. C. (1979) Interestingness: Controlling inferences. *Artificial Intelligence* 12:273–97. [SJR]
- Schiffer, S. (1972) *Meaning*. Clarendon Press. [aDS]
- Searle, J. (1969) *Speech acts*. Cambridge University Press. [aDS]
- (1979) Metaphor. In: *Metaphor and thought*, ed. A. Ortony. Cambridge University Press. [MD]
- (1983) *Intentionality*. Cambridge University Press. [aDS]
- Seuren, P. A. M. (1988) The self-styling of relevance theory. *Journal of Semantics*, 5:123–43. [rDS, PAMS]
- Simon, H. A. (1979) *Models of thought*. Yale University Press. [JEA]
- Slobin, D. & Aksu, A. (1982) Tense, aspect, and modality in the use of the Turkish evidential. In: *Tense-aspect: Between semantics and pragmatics*, ed. P. J. Hopper. John Benjamins. [NVS]
- Smith, N. V.; ed. (1982) *Mutual knowledge*. Academic Press. [aDS]
- (1987) Relevance theory and first language acquisition. In: *Encontro de linguistas: Actas*, ed. H. Alves. Universidade do Minho. [NVS]
- Smith, N. V. & Smith, A. (1988) A relevance-theoretic account of conditionals. In: *Language, speech and mind: Studies in honor of Victoria Fromkin*, ed. L. Hyman & C. Li. Croom Helm. [NVS]
- Sperber, D. (1975) Rudiments de rhétorique cognitive. *Poétique: Revue de Théorie et d'Analyse Littéraire* 23:389–415. [rDS]
- (1984) Verbal irony: Pretense or echoic mention? *Journal of Experimental Psychology: General* 113(1):130–36. [aDS]
- (1985) *On anthropological knowledge*. Cambridge University Press. [aDS]
- Sperber, D. & Wilson, D. (1981) Irony and the use-mention distinction. In: *Radical pragmatics*, ed. P. Cole. Academic Press. [aDS]
- (1982) Mutual knowledge and relevance in theories of comprehension. In: *Mutual knowledge*, ed. N. Smith. Academic Press. [rDS]
- (1986a) *Relevance: Communication and cognition*. Harvard University Press and Blackwell.
- (1986b) Loose talk. *Proceedings of the Aristotelian Society (New Series)* 1985/6:153–71. [arDS, MD]
- Stalnaker, R. (1974) Pragmatic presupposition. In: *Semantics and philosophy: Studies in contemporary philosophy*, ed. M. Munitz & P. Unger. New York University Press. [DB]
- Strawson, P. (1964) Intention and convention in speech acts. *Philosophical Review* 73:439–60. Reprinted (1971) in *The philosophy of language*, ed. J. Searle. Oxford University Press. [aDS]
- (1969) Meaning and truth. Inaugural lecture at the University of Oxford. Reprinted (1971) in *Logico-Linguistic papers*. [aDS]
- (1971) *Logico-linguistic papers*. Methuen. [aDS]
- (1980) Reply to McDowell. In: *Philosophical subjects*, ed. Z. van Straaten. Oxford University Press. [MD]
- Taglicht, J. (1984) *Message and emphasis: On focus and scope in English*. (Longman. [aDS]
- Waltz, D. L. (1987) Connectionist models: Not just a notational variant, not a panacea. In: *Theoretical issues in information processing 3: Position papers*, ed. Y. Wilks. Computing Research Laboratory, New Mexico State University. [EH]

## References / Sperber & Wilson: Relevance

- Wilks, Y. (1986) Relevance and beliefs. In: *Reasoning and discourse processes*, ed. T. Myers, K. Brown & B. McConigle. Academic Press. [PAMS]
- Wilks, Y. & Ballim, A. (1987) Multiple agents and the heuristic ascription of belief. Proceedings of the International Joint Conference on Artificial Intelligence, Milan. [YW]
- Wilks, Y. & Cunningham, C. (1986) A purported theory of relevance. In: *Language and protest: A festschrift for Petr Sgall*, ed. J. Mey. John Benjamins. [YW]
- Williams, E. (1981) On the notions "lexically related" and "head of a word." *Linguistic Inquiry* 12:245-74. [RK]
- (1986) A reassignment of the functions of LF. *Linguistic Inquiry* 17(2):265-300. [RK]
- Wilson, D. & Sperber, D. (1981) On Grice's theory of conversation. In: *Conversation and discourse*, ed. P. Werth. Croom Helm. [JLM]
- (1986a) Inference and implicature. In: *Meaning and interpretation*, ed. C. Travis. Oxford University Press. [rDS]
- (1986b) Pragmatics and modularity. In: *Chicago Linguistic Society 22: Parasession on pragmatics and grammatical theory*. [aDS]
- (1988a) Mood and the analysis of non-declarative sentences. In: *Human agency: Language, duty and value*, ed. J. Dancy, J. Moravcsik & C. Taylor. Standord University Press. [NVS]
- (1988b) The self-appointment of Seuren as censor. *Journal of Semantics* 5:148-62 [PAMS]
- Wright, R. (1975) *Meaning-*nn* and conversational implicature*. In: *Syntax and semantics 3: Speech acts*, ed. P. Cole and J. Morgan. Academic Press. [aDS]
- Yu, P. (1979) On the Gricean program about meaning. *Linguistics and Philosophy* 3(2):273-88. [aDS]
- Ziff, P. (1967) On H. P. Grice's account of meaning. *Analysis* 28:1-8. Reprinted (1971) in *Semantics: An interdisciplinary reader*, ed. D. Steinberg & L. Jakobovits. Cambridge University Press. [aDS]
- Zubizarreta, I.-L. (1985) The relation between morphophonology and morphosyntax: The case of romance causatives. *Linguistic Inquiry* 16(2):247-89. [LH]

### POSITION AVAILABLE

OSU's Laboratory for Artificial Intelligence invites people with strong backgrounds in cognitive science, a PhD, and advanced training in computer science to apply for faculty positions in the Department of Computer and Information Sciences. Areas of emphasis will be computational theory issues surrounding AI models including connectionist schemes, computational approaches to vision and speech understanding, and knowledge-based problem solving. Cognitive science studies, with participation by computer and information science, linguistics and cognitive psychology programs, among others, are to be supported; a senior appointment as Director of Cognitive Science is possible.

Candidates should send a vita, reprints, a description of backgrounds and interests, and names of at least three references to:

Dr. B. Chandrasekaran  
Director of LAIR  
Dept. of Comp. and Inf. Sciences  
2036 Neil Avenue Mall  
Columbus, OH 43210-1277

Women and minorities are especially encouraged to apply. The Ohio State University is an Affirmative Action/Equal Opportunity Employer.