J. Linguistics 25 (1989), 455–472. Printed in Great Britain

REVIEW ARTICLE

A review of Relevance

STEPHEN C. LEVINSON

Department of Linguistics, University of Cambridge

(Received 31 March 1989)

Dan Sperber & Deirdre Wilson, *Relevance: communication and cognition.* Oxford: Blackwell, 1986. Pp. viii + 279.

In this bold and highly controversial book, Sperber and Wilson attempt to shift the whole centre of gravity of pragmatic theory by locating it firmly in a general theory of cognition. Outlining that general theory takes the bulk of the book, so those who have followed technical developments in pragmatics will not find those issues much advanced here. For the purpose of the book is otherwise, to outline a single cognitive principle which (it is held) will give us, along with a theory of attention, almost all we need in the way of a theory of communication, such a theory in turn having linguistic applications merely sketched here.[1] The book is written in a fluid argumentative style; easy to read, it is not easy to understand, presuming much that the central thesis depends on. Perhaps because of the global and speculative aims, there is little or no reference to recent developments in the theory of meaning.[2]

This book has already aroused much passion for and much passion against. Cited by its authors as long ago as 1979, the ideas have been trailered in a series of articles, with both critics and protagonists taking up provocative positions (see for example the still useful exchange in Smith, 1982). Thus when the book finally appeared, extensive airings of these passions were inevitable (see the peer review in *Behavioral and Brain Sciences* (*BBS*) **10** (1987), or the exchange in *Journal of Semantics* **5** (1988)).[3] The views have

[1] A second and third volume on pragmatics and rhetoric were projected (vii–viii) and this reviewer had hoped for some glimpse of these applications before passing judgment on a disembodied principle, as it were. But there are now a number of applications of these ideas to linguistic detail, and these perhaps give a better idea of the linguistic interest of the theory than anything in the book itself; see e.g. Kempson, 1986; Carston, 1988, on the semantics/pragmatics interface; Blakemore, 1987, on conventional implicatures and discourse particles; Kempson, 1988, on anaphora; Smith, 1981; Haegeman, this volume, on tense; Smith & Smith 1988, on conditionals.

[2] For example, there is scarcely any mention of recent work on generalized implicature (see below) or on the new contextually-sensitive semantic theories or on computational pragmatics or on computational theories of inference. There are no bibliographic references to, among others, Atlas, Horn, Barwise & Perry, Heim, and Kamp, despite the relevance of the work of all those authors to the themes discussed.

[3] The extended history of the ideas, and battles over them, perhaps helps to explain the often vituperative nature of the controversy.

been energetically promoted by the authors (who have published numerous précis elsewhere) and by other adherents. For instance, the ideas play a central role in three of the survey articles in Newmeyer's (1988) *Linguistics: the Cambridge survey*. Thus the views cannot be ignored.

This reviewer might as well declare from the outset that, as regards the central thesis of the book, he belongs with the 'nays' and, he believes, the nays must have it. The reasons for the negative assessment of the central thesis lie on many levels; the book relies on improbable presuppositions about human cognition; it underplays the role of usage in pragmatic theory; it ignores many current developments in semantics, pragmatics and the study of inference; it is too ambitious and globally reductive; and anyway the theory is obscure and it is not clear how it could be made to have clear empirical application. But even presuming that this negative assessment of the thesis is correct, the book is important for a number of reasons: it draws central attention to the role of contextual inference not only in language comprehension but in what many have taken to be the heart of semantics; it has interesting things to say about the nature of context; and it constitutes a bold attempt to rethink pragmatic theory from the base, considering how it should fit in with theories of cognition. We are offered, as an alternative to the admittedly ramshackle but empirically-based observations of much contemporary pragmatics, a simple unifying vision – foxes and hedgehogs, splitters and lumpers, cautious clerics and academic prophets could hardly be more opposed. Thus, regardless of the fate of their thesis, we owe Sperber and Wilson a debt for bringing all these issues to the forefront.

Yet this reviewer is, like any working pragmaticist, an interested party, and such is the status of the book, as *cause célèbre* in the theory of communication, that he can only urge readers to take a look at it and come to their own opinion. In this review we can merely indicate the nature of the central thesis, and help to crystallize the issues around which discussion might most usefully revolve.

1. Brief outline

The book has four chapters. In the first ('Communication') the origins of the whole theory become clear. In the early 70s Grice's account of meaning$_{nn}$ was thought by many to give us for the first time the basis for a general theory of communication (see for instance Lewis, 1969; Schiffer, 1972); however, various outstanding problems were left beached as philosophy of language began rapidly to go out of fashion. Of these, one concerned the notion of mutual recognition of intentions (and how this was to be achieved without infinite regress of knowledge states) and another concerned how it is that the content to be communicated could have the natural saliency such that sender and receiver could co-ordinate on its production and recognition. Sperber and Wilson's book is an attempt to answer these questions, and then to show

that these answers provide more general insights. The answers provided in Chapter 1 are: (a) the problems associated with a notion of mutual knowledge can be avoided by substituting a weaker concept of mutual manifestness;[4] (b) the requisite notion of saliency can be provided by a theory of cognitive optimization, to be dubbed Relevance, and not (as the theorists of the 70s had supposed) by a theory of interactive co-operative co-ordination (whether as sketched by Lewis (1969) or Grice's Co-operative Principle). The same notion of saliency can then be used in the definition of 'mutually manifest' in (a).

Chapter 2 outlines a speculative psychology of inference likely to leave psychologists, logicians, semanticists and computer scientists in some degree of apoplexy.[5] All inference, it is claimed, is deductive (even if the premises have to be guessed in various ways) and proof-theoretic, that is to say computed over the syntax of logical forms. Sperber and Wilson make a large number of very specific assumptions here and they offer, with charming hubris, the 'outline of a formal deduction system intended to model the system used by human beings in spontaneous inference' (94). There have been various attempts (unreferenced by Sperber and Wilson) to restrict deductive processes so that they are more in line with psychological plausibility; Sperber and Wilson opt for a simple elimination of 'introduction-rules' (e.g. the inference from *p* to *p or q*) in order that the implications of any deduction will be strictly finite (a move required to make sense of their principle of Relevance).[6]

This chapter has to be read in the context of a modified Fodoreanism, that is an assumption of cognitive modularity, with a distinction between 'central processor' and specialized input modules (vision, language, etc.).[7] The modification is that whereas Fodor assumes the 'central processor' works mysteriously along non-deductive lines (and is thus sceptical of scientific advance here), Sperber and Wilson present their deductive system as the central processor itself, processing logical forms provided by the different input devices. The logical forms are 'structured sets of concepts', each concept associated with an address in memory to which are attached further concepts that may be processed. A freshly input logical form can be

[4] The alleged advantages of mutual manifestness over mutual knowledge have been a focal point in discussions of Sperber and Wilson's ideas, and they need not be discussed here; see Smith (1982) and the *BBS* peer review.

[5] See for example the articles by Gazdar and Good, and by Wilks in Smith (1982); and the comments by Hinkelman, Macnamara, Russell and Seuren in *BBS*.

[6] No proofs or formal properties of the system Sperber and Wilson envisage are actually given. It is not self-evident for example that the removal of introduction-rules, but the retention of commutativity laws (*BBS* 10: 741), will be sufficient to ensure finite consequences from finite premises (see Seuren's comments, *BBS* 10: 733). Besides, as Gazdar and Good (1982: 90) pointed out, some sets of intuitively non-trivial inferences would seem to be essentially infinite.

[7] See Fodor (1983), whose ideas and style form an essential backdrop to this book.

processed to yield its intrinsic deductive content, or it can be processed together with existing assumptions to yield CONTEXTUAL IMPLICATIONS (the inferences that follow only from the conjunction of the old and new information).

This chapter is out of line with nearly all the received wisdom about the nature of intuitive and ampliative inference, views which have motivated the many twentieth-century experiments with non-deductive reasoning systems (with inductive calculi, mental models, abductive and practical reasoning systems, and most recently non-mononotic default systems).[8] Sperber and Wilson's justification for the departure is, essentially, the parsimony of bold simplicity. But there is a narrow line between the simple and simplistic, here surely transgressed. And the supposedly simple turns out to have many difficulties of its own.[9]

---

[8] Of these, they discuss only inductive calculi, arguing that degrees of inductive strength can be captured in terms of the processing history of assumptions; Johnson-Laird's (1983) 'mental models' approach which argues that all the experimental results point to deduction-like results being achieved by non-deductive reasoning, is rejected without argument except that deduction is efficient (102).

[9] It would take a much longer review to spell all these out; but see for example Seuren's *BBS* commentary (any critique is hampered by the fact that, as mentioned, no formal properties of such a system are given). One example is the attempt to justify the absence of an &-introduction rule. In order to handle the inference:

$$\frac{(p \& q) \rightarrow r \quad p \quad q}{r}$$

the authors include the following rule in the logical entry (in the memory location) for '&':

$$\frac{(p \& q) \rightarrow r \quad p}{q \rightarrow r}$$

This together with:

$$\frac{q \rightarrow r \quad q}{r}$$

will get us the same result, namely *r*. But how then are we to get the corresponding inference (cf.their example on p. 109):

$$\frac{(p \& q \& s) \rightarrow r \quad p \quad q}{s \rightarrow r}$$

Avoiding &-introduction would seem to lead to an infinity of elimination rules. (Unless, of course, the sentential variables here are schematic rather than atomic letters; but if they are schematic we would need a complex parser that never lost track of the atomic propositions.) Faced with these sorts of problems, Sperber & Wilson in the end retreat and admit the use of non-deductive heuristics to ape deduction (102). The admission is damaging: any experimental results favouring 'introduction-rule' reasoning can be attributed to these heuristics, making their claims untestable; and once one has admitted heuristics-that-ape-deduction, why not go the whole hog as Johnson-Laird (1983) advises?

Chapter 3, 'Relevance' (R), brings us to the heart of the matter. This is the cognitive principle which is to give us at once (a) the account of saliency so needed in a Gricean theory of communication, (b) the principle controlling the operations of the deductive device. It amounts, disappointingly perhaps, to just the 'biggest bang for the buck', here essentially maximal deductive yield for minimal processing effort. Thus, on the analogy to industrial productivity (125), R equals output over input, or:

$$R = \frac{\text{E (number of contextual effects)}}{\text{C (cost of effort involved in obtaining E)}}.$$

We are warned though that there may be no single currency in which to measure E and C, for C may vary with one's mood, tiredness, etc., and be assessed in terms of 'symptomatic physico-chemical changes' (130).[10] Indeed, R is not itself mentally represented or an object of computation, but is more an extrinsic constraint on the processing system (132).

Nevertheless, if R is to be objectively assessed, there must be some measures of E and C we can turn to. The measure of processing cost C is admittedly problematic (130). Without justification, Sperber and Wilson later seem to presume that there is some relation between the surface complexity of English syntactic structures and the effort required to process the logical forms derived from them (by the linguistic module, not the central processor). E, on the other hand, can be characterized as the sum of contextual effects, consisting of: (i) the contextual implications of a logical form in the context of other assumptions (and we now see why we must limit deduction to obtain finite implications); (ii) the strengthening or weakening of old assumptions.[11] Since the yield of contextual implications itself depends on the size and content of the 'context' (understood as the set of salient accessible assumptions), E will be unrestricted in value unless the context is somehow strictly limited. How? Well by R of course! The context is expanded from that minimally given by the last state of the deductive device, by incrementally looking up further connected assumptions in memory, and so on, until processing achieves a set of contextual effects balancing the effort required to derive them. R thus controls the basis for assessing R; this is one of a number of apparent circularities in the chapter (see below).

All this has nothing special to do with language. R applies equally to the

[10] Sperber and Wilson's caution here is meant to invalidate the well-known critique by Gazdar and Good (1982) of any kind of numerical equation of this sort. See footnote 15 below.

[11] Having rejected a calculus of subjective probabilities, Sperber and Wilson here presume some computation of rough confirmation values (111 ff.). They hope thereby to enable the deductive device to handle contradictions without generating infinite conclusions, by accepting the premise with the higher confirmation value. Sperber and Wilson have to add confirmation values to their account of Relevance, otherwise reminders, non-obvious tautologies, and the like would be ir-Relevant.

processing of smells, noises, trains of thought and the whole structure of attention, for R is a globally ambitious theory of central thought processes (151 ff.). But equally R must apply to linguistic input. Indeed, communication turns out (it is claimed) to rely critically on R, as a communicative act works by presenting a stimulus that is naturally salient given R and is presented as worth processing by the standard of R. We are then given (as the last of a series of definitions) the PRINCIPLE OF RELEVANCE: 'every act of ostensive communication communicates the presumption of its own optimal relevance' (158). This principle is explicitly intended to replace Grice's Co-operative Principle with its well-known maxims of conversation, and the differences are listed (161–162): (i) R is weaker but more explicit; (ii) R is an automatic principle that works, unlike the CP, without any overt knowledge of it; (iii) R applies both to 'what is said' and 'what is implicated'.

It is only in the last 70 pages, Chapter 4, that linguistic issues are directly addressed. The Fodorean dogma of grammar as an input system, and the Chomskyan dogma that grammar and communication are only accidentally linked, are the opening premises. The inferential basis of communication (à la meaning_{nn}) is, it is claimed, independent of the subservient specialism of coded verbal communication, which always ultimately depends on the inferential umbrella (a general sentiment in pragmatic theory but forcefully expressed here). The goal of pragmatics, we are told, is to explain how a 'hearer' (*sic*) identifies 'which assumptions made manifest by [the speaker's] utterance are such that it is mutually manifest that the speaker intended to make them manifest' (179). The task can be broken down into a number of subtasks: (a) disambiguation, (b) identification of the proposition expressed (for example by resolution of anaphora), (c) identification of the intended mood, (d) the calculation of Relevant contextual implications.

The most interesting ideas here are the very brief discussions of (a) and (b). Grice noted that identification of 'what is said' depends on prior disambiguation; Sperber and Wilson claim that R is crucially involved – the chosen reading is the one that gives most contextual effects (for least effort as always). The authors recognize a problem here: if grammar is a Fodorean input module, there would be no feedback possible from the central processor to the control of the processing in the input module, so every reading of some constituent would have to be submitted to the central processor (where R rules supreme) for disambiguation by R (186). Turning to the identification of the proposition expressed, Sperber and Wilson note that sentences do not generally express complete propositions: 'semantic representations are incomplete logical forms, i.e. at best fragmentary representations of thoughts' (193). A semantic representation has thus to be EXPLICATED into a full logical form by the resolution of anaphora, and the narrowing of semantic generality (thus *John's book is good* might express the proposition 'the book by John is good' or 'the book John is reading is good', etc.). Such a 'development' of a logical form is called an EXPLICATURE and

is to be distinguished from an IMPLICATURE, which is a contextual implication. Yet the notion of explicature is never properly defined nor satisfactory diagnostics given. What guides the development of an explicature? R of course; whatever fleshing out of a logical form maximizes deductions (or at least contextual effects) will be favoured. The notion of an explicature has an implicit anti-Fodorean corollary: semantics, understood as the assignment of complete logical forms and their truth-conditions, can no longer belong to the language input module, which will now output mere fragments of logical form.[12]

Implicatures are now just contextual deductions. Sometimes though, as in incomplete answers, one may need to hypothesize an extra premise in order to get a deduction that would make an utterance Relevant:

A: 'Would you fly a Zanussi?'
B: 'I wouldn't fly any cheap helicopter'
implicated premise: 'A Zanussi is a cheap helicopter'
implicated conclusion: 'B wouldn't fly a Zanussi'

Grice's examples of Relevance implicatures are indeed like these implicated premises: but Sperber and Wilson do not dwell on the fact that they are formed not by deduction but by the kind of creative hypothesis formation that most observers have thought to be central to a theory of implicature (see below). The authors do have some interesting things to say about the INDETERMINACY of some implicatures; there may be a number of further potentially implicated premises (for example 'cheap helicopters are dangerous'), that may yield further deductions ('Zanussis are dangerous'), and there may be doubt as to whether they were meant to be mutually manifest (196 ff.).

The chapter contains sections dealing with the pragmatic effects of stress and marked word order, presupposition, tropes and speech acts, each too short to be more than a summary of the authors' previously published views on these subjects, without defence against earlier objections. Each is likely to raise more questions in the reader's mind than it answers; for instance how the central processor has access to surface features like stress and word order, how the treatment of presuppositions as entailments could possibly explain their defeasibility and projection properties, and so on. The book ends abruptly without conclusions.

2. ASSESSMENT

As should be clear from the summary, this book is more a polemical attempt to force a paradigm shift in favour of a particular approach to cognition than

[12] This corollary of Relevance theory is explored in Kempson, 1986, and is discussed further below. One unexplained problem with this picture is how the central processor can reason at all with un-explicated fragmentary logical forms as input.

it is a detailed contribution to linguistic pragmatics.[13] The approach requires the reader to buy a set of very specific, and to many minds implausible, assumptions: Fodorean modularity of a specific kind, with a restricted kind of deduction as the heart of central processing, a translation of all input perception into a simple canonical 'logical form', a governing cognitive principle that is not itself an object of processing, a subjective measure of effort being the main limit on processing, and so on. If you do not buy these assumptions, the theory will be untenable in its entirety. Even if we buy these assumptions, the theory to my mind never quite gets off the ground for two reasons: the first is lack of internal coherence and clarity, and the second uncertainty of external application. Let me take these up in turn.

A central puzzle is exactly how R is really meant to work. Let us take for example the claim that R predicts a unique, intuitively correct, interpretation of an utterance.[14] As outlined, R is a function of (measures of) contextual effects (E) balanced by processing costs (C), related, say, as R = E/C (though not thus subjectively assessed).[15] This allegedly yields unique interpretations. Yet a close reading of the text indicates that how the equation is solved varies. Sometimes R has a predetermined value V, such that contexts are

[13] The polemical nature of the writing has produced more heat than light (see for example the exchange in *BBS*). For example, neo-Gricean pragmaticists are treated as strawmen, allegedly reducing pragmatic inference to coded communication; in fact there is a large measure of agreement on the very features that Sperber and Wilson think so distinctive of their own work, like the essentially inferential nature of pragmatics (see for example the responses by Bach & Harnish and Recanati in *BBS* 10:712). There is a lack of proper, helpful reference at many points; for example, in a theory aiming to give an account of implicature it is distinctly odd to have not a single reference to the work of Atlas, Horn, Hirschberg, and others; it is also strange to have no reference to earlier attempts to construct a concept of relevance (e.g. Dascal, 1977); and, given the computational nature of the theory, peculiar not to have extensive reference and comparison to problem-solving and inference in AI. There is inconsistent hyperbole, so that, for example, we are sometimes told that R will pick out exactly one interpretation (see next footnote), sometimes that interpretations will depend on how tired a 'hearer' is (131–132), that they may be indeterminate, and so on. Stipulation often replaces proper argumentation, as with the dismissal of Johnson-Laird's ideas (102). These would be serious complaints against a work in a less polemical or speculative genre; but they do render the book unusable as a text, and potentially hazardous in the hands of the untutored.

[14] The power of R to yield unique determinate solutions (169) is further stressed in Wilson & Sperber (1986: 80): 'the first accessible enrichment consistent with the principle of R is the *only* enrichment consistent with the principle of R'; 'it is easy to show that every utterance has at most one interpretation which is consistent with the principle of relevance' (1986: 76).

[15] Gazdar and Good (1982) discussed at length the difficulties of taking R to be a function of this sort. Sperber and Wilson (1982b: 109 n. 1; and in the book under review, p. 130) in turn reject the characterization, but what they suggest instead only increases the puzzle – R must achieve (at least approximately) the effects of such a function without actually being computed in this way. See also Hinkelman (*BBS* 10:720). For the record, Hinkelman and Levinson are said by Sperber and Wilson to confuse maximal relevance and optimal processing for optimal relevance (*BBS* 10: 744–745). It remains unclear why optimal R is supposed to escape these dilemmas.

expanded till V is satisfied (142) (but since context can be expanded in many different, equally R-full ways, how can we get the one correct expansion?). Sometimes R is a comparative measure, selecting the best interpretation from competing ones (144, 153). This would clearly involve comparison of different interpretations; but obviously to compare interpretations would involve computing each of them, thus additively increasing C, so R predicts it will be cheaper to pick any one interpretation at random! (Sperber and Wilson are aware of this difficulty, and suggest heuristics must exist that allow estimation of costs without computation (131); and so they must, and must be described and shown to work, if the theory of R is to survive.) Sometimes, apparently, C has a theshold value such that the first accessible E-yielding context is automatically selected (178, 185; see also my footnote 14). But if only R is fixed, there is no determinate solution; if C is fixed too, then comparing interpretations might just exhaust C without any solution. And if R and E are fixed, then again we may fail to find either a unique solution or any solution yielding sufficient E for the computed C. In fact none of these ways of solving the equation are guaranteed (or even likely) to give a unique solution; and each way of solving the equation is likely to give a different solution. The reader thus comes away from each sample analysis with the distinct impression of sleight of hand, not the impression of a demonstration that R can explain the inferences in question.[16]

The second main problem derives directly from the first. If the account is to be of any use, R must make clear predictions, or at the very least it should be possible to falsify R. But unfortunately R doesn't seem to have that kind of clear application. For one thing, in the present state of psycholinguistic knowledge, we have no way of estimating processing costs C (Sperber and Wilson, no doubt correctly, reject some operational estimate in terms of number of deductive steps (130)). Secondly, it is always possible to adjust the alleged contextual premises to make the account fit the scheme. Thirdly, as described, R could be rather differently computed with rather different results. The mode of argumentation lends further to this sense that the theory is self-fulfilling: the job is to show that the intuitively correct interpretation of some utterance conforms with R, indeed is the only interpretation that conforms to R; we then try to establish how it so conforms, finding a good enough *post facto* story about large amounts of effects E in some hypothesized context for small amounts of effort C. But this simply doesn't look like any

[16] Clearly this reviewer is not alone in this impression: compare for example: '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' (Russell's comment in *BBS* 10:731). Other commentators express an exasperation with the special mix of bold assertion and lack of clarity that pervade the book.

advance on the *ex post facto* Gricean accounts that Sperber and Wilson complain about (37).[17]

However, above and beyond all this, what *a priori* reason is there to think that these two factors E and C are the crucial factors determinative of utterance interpretation? E is a very specific measure of informational richness; most pragmatic theorists assume two such measures, incommensurable, indeed intrinsically opposed to each other (see for example Atlas & Levinson, 1981; Horn, 1984, both drawing on many sources). C is a measure of inference-cost; yet all the evidence points to inference being 'cheap' and profligately employed without waiting for input that would make it unnecessary, while a quite different kind of cost, associated with the production of marked and prolix expressions, is demonstrably determinative of language use (as shown for instance by Zipf, 1949). Sperber and Wilson rightly insist on the fundamental role of inference in coded communication; but the very reason for that role is the extraordinary effortlessness of human inference.

Sperber and Wilson write with the fervour of those who have discovered the philosopher's stone, a single principle that will render tractable all the mysteries of attention, trains of thought and, as a mere by-product as it were, language comprehension. Many observers simply cannot understand why they think there should be such a thing, let alone why they think R, with its uncertainties of construction and application, could be it.

For the non-believers, there is still much of value in this book. There is a serious attempt to patch up Grice's theory of meaning$_{nn}$; and this should be compared to the rather different tack taken for example by Perrault and Cohen in AI (see Perrault, 1987). There is a forceful emphasis on the essential role of inference in the interpretation of coded communication. This is of course a tenet shared by all pragmaticists (*pace* Sperber and Wilson, pp. 14 ff.), but the authors push the account further and in most interesting ways, arguing that disambiguation and propositional determination depend fundamentally on pragmatic inference (more on this in a moment). There is an important argument against the idealization wherein context is taken to be the conversational 'common ground', and in favour of a view where context is seen as a set of premises invoked from that background (or constructed if necessary) by pragmatic principles.[18] There is an attempt,

[17] Sperber and Wilson's R is supposed to be a simple computational model, and so it is reasonable to ask that it be clarified by a flow-chart. Attempts by readers to do this founder on just too many uncertainties (see for instance Hinkelman, in *BBS* 10:721; my own attempt came out rather differently). Sperber and Wilson would do us a service by trying themselves. One problem here is how to compute incommensurables -- effort is metered in 'physico-chemical changes', while effects are measured in two currencies (i) the cardinality of contextual implications; (ii) confirmation strengths.

[18] The invocational view of context is one currently being explored in a number of rather different traditions: for example, in AI under the rubrics of 'frames' and 'focus' (see Sidner, 1986), and in anthropological linguistics under the rubric of 'contextualization' (see Gumperz, 1983; Silverstein, 1988).

perhaps for the first time, to say something systematic about the indeterminacy of pragmatic inferences (a thesis not sufficiently pursued in the book). Above all, and most controversially perhaps, there is a forceful attempt to shift the whole centre of gravity of pragmatic theorizing away from the study of usage principles to the study of cognitive principles.

3. Some leading issues

The theory of meaning has been dogged by the tendency to generalize from the nature of meaning in one limited linguistic phenomenon to the nature of meaning as a whole. For Sperber and Wilson the focal phenomena are the classic particularized Relevance implicatures illustrated by Grice, typified by the partial answer to a question. Meanwhile, the implicatures which have been focal to Neo-Gricean pragmatic theory (and to Grice's own central interest with the pragmatics of the connectives) are the generalized conversational implicatures (GCIs) attributable to the two maxims of Quantity and to Manner (see Horn, 1972; Gazdar, 1979; Atlas & Levinson, 1981; Horn, 1983, 1984, 1985; Levinson, 1987a, b; Atlas, 1989; Horn, 1989). These latter inferences are scarcely mentioned in the book, except to dismiss them as 'untypical examples' (36–37), best treated as explicatures (262 fn 4).[19] They have been focal to the Neo-Griceans because they have immensely regular, cross-linguistic generality, and interact closely with linguistic structure and meaning. For example, on the basis of GCI theory we are able to predict, thanks to Horn (1972, 1989), that no language is likely to lexicalize 'not all' in the same way that it is likely to lexicalize 'not some' as say *none*. Cross-linguistic predictions of that sharpness are actually rather rare in linguistic theory, and are not to be lightly dismissed.

Sperber and Wilson would no doubt agree that GCIs and the classic Relevance implicatures are different in kind (the former are presumably always explicatures, although that notion is never properly defined); but they are committed to the wholesale reduction of implicature (including GCIs) and the whole Gricean apparatus of maxims to R (161).[20] Apart from the impossibility of reducing countervailing principles to one mega-principle (see

[19] Sperber and Wilson's response to my earlier complaint about this dismissal is to repeat the charge and to assert that 'His [Grice's] best-known examples are particularized implicatures; the discussion of generalized implicatures is restricted to a few cases; and there is no evidence that he [Grice] saw the distinction as theoretically significant' (*BBS* 10: 748). In fact, Grice's main purpose in developing the theory of implicature was to give an account of the meaning of natural language connectives in terms of truth-functional content overlaid by generalized conversational implicature. Hence the bulk of Grice's unpublished William James lectures is devoted to GCIs.

[20] Curiously, they don't repeat in this book their earlier (Wilson & Sperber, 1981) detailed attempt at reductionism of all the Gricean maxims to R, apparently taking it as more or less self-evident. (For arguments against the reduction see Horn, 1984; Levinson, 1987b; Sadock, 1986.)

for instance Horn, 1984, 1989; Levinson, 1987b; Sadock, 1986), the central problem with this reductionism is that GCIs are GENERALIZED, that is they go through unless there are reasons to think otherwise; in short they are preferred interpretations, and fall into the category of non-monotonic default inferences now being energetically explored in Cognitive Science (see Ginsberg, 1987). But R cannot predict any stable kind of inference without a constant context, since it is the derivation of particularized contextual effects that satisfies R. R is thus a theory of 'nonce-inference', of speaker-meaning, and utterance-token-meaning. GCI theory is in contrast a theory of preferred or default interpretation, of utterance-type-meaning. There can be no way to get the universal regularities of GCI predictions out of a theory of 'nonce-inference'; relevance theory was simply not concocted to deal with GCIs, and should relinquish ambitions in that direction. (Equally, GCI theory may have little to say about 'nonce-inference', although see Hirschberg's (1985) unification of generalized and particularized Quantity implicatures.)

The question then arises whether R at least offers a reasonable account of the classic Relevance implicatures which are the focal phenomena that the theory was designed around. Curiously, the account seems strangely inadequate here, even assuming that the internal difficulties with the theory could be patched up. The problem is that the essentially interesting features of these inferences are non-deductive. Consider the following (or equally Sperber and Wilson's example (5) on pp. 121–122):

A: 'Where are my chocolates?'
B: 'The dog is looking very sleepy.'

Somehow, presuming the relevance (or co-operativeness) of B's reply, we supply not only an implicated conclusion (perhaps 'the dog may have eaten your chocolates') but also the necessary premises (say, 'if dogs eat chocolates they get very sleepy').[21] The premises may even be counterfactual (as here, or premise (e) in Sperber and Wilson's more complex example just cited), so they cannot just be retrieved from memory as the authors sometimes seem to suggest. So here lies the central mystery posed by this kind of implicature: WHAT KIND OF INFERENCE IS SO POWERFUL THAT IT CAN PROVIDE BOTH THE PREMISES AND THE CONCLUSIONS to an argument? (Levinson, 1985). Unfortunately, the theory of R does not help; it merely tells one how to check the R-value of the conclusion once you have the premises. Sperber and Wilson recognize the role of such creative hypothesis formation, but they somehow lose sight of its centrality, claiming instead that 'the formation of

[21] Of course this will not be a premise in a deductive argument, but it might be one in an abductive one (see Hobbs, 1987).

assumptions by deduction is the key process in non-demonstrative inference' (83).[22]

A further problem is that the pre-theoretical notion of relevance is not well captured by R, because R is an informational measure (balanced by effort), whereas pre-theoretical relevance is largely about the satisfaction of others' goals in interaction, and the satisfaction of topical and sequencing constraints in discourse, as in the expectation that an answer will follow a question in the example above (see for example Dascal, 1977; Holdcroft, 1987; Levinson, 1987b: 74 ff.). Although Sperber and Wilson respond by saying (119 f.), in effect, so much the worse for pre-theoretical relevance, their R will leave quite unexplained the tie between questions and answers, between questions and helpful non-answers (A: 'What's the torque wrench look like', B: 'There's one in front of you'), between greetings and greetings, and in general all the things some theory of relevance ought to explain.[23]

Thus, on the face of it, R fails to account both for the type of implicature focal to GCI theory and for the prototype cases of implicature attributable to pre-theoretical relevance. But Sperber and Wilson are of course correct that we do need theories of particularized implicature. Here, readers should know that there are alternative theories available, more closely attuned to concepts of pre-theoretical relevance. There are a number of interesting attempts to model the inference to others' interactional goals within the literature on AI planning systems (Allen, 1983; Pollack, 1986), and these have been made the basis of a theory of implicature by Thomason (1987) and McCafferty (1987). Discourse analysts, of course, have attempted to spell out the sequencing and topical constraints underlying the notion of a pertinent response (see e.g. Brown & Yule, 1983; Levinson, 1983: Ch. 6). Meanwhile, there are a number of interesting computational theories of 'nonce-inference' (often purveyed under the rubric of LOCAL PRAGMATICS), including Hobbs' (1987) abductive account of implicature, and the outstanding work by Hirschberg (1985; in press) on non-generalized Quantity implicatures.

One final issue: as we have noted, Sperber and Wilson draw attention (under the rubric of 'explicatures') to the role of pragmatic inference in the

[22] In reply to an earlier complaint of mine about this problem, Sperber and Wilson (*BBS* 10:749) gesture to heuristics, like affirming the consequent. But that of course is exactly the focus of the theories of abductive inference which they eschew.

[23] Another way of putting the problem: pre-theoretical relevance is an INTERACTIONAL issue – A's utterance is relevant for a particular interlocutor B trying to do something in particular. In contrast R is a cognition-internal measure; what is to guarantee that R-for-A is R-for-B (the Principle of R, in effect 'say something which is optimally R', does not specify R-for-whom)? Pre-theoretical relevance is partly about bridging A's concerns and B's concerns, about A taking cognizance of B's point of view, etc. (cf. Millikan's comment in *BBS* 10:725f.). Incidentally, Sperber and Wilson's own position is none too clear; on the one hand R doesn't need to be known, let alone computed, because it is a mental automatism (162); on the other mutual awareness of the Principle of R seems an essential prerequisite to utterance processing.

determination of a proposition expressed by an utterance. This is one of the most important issues raised in the book, but this potentially controversial point is nicely illustrated rather than systematically argued for. Select philosophers have long held this point of view (Strawson, 1952), while Atlas (unreferenced in the book) has attempted for many years to push formal semanticists to swallow this bitter pill. But beginning in the early 1980s a slow realization of the extent of contextual determination of propositional content had begun to dawn on all concerned with the theory of meaning. AI workers dubbed the process of propositional determination 'local pragmatics' (Hobbs & Martin, 1987; Pereira & Pollack, 1988) and GCI theorists (Atlas, 1979; Levinson, 1987b) had independently or by osmosis come to similar conclusions. Formal semantics too underwent drastic modifications in different directions in the hope of dealing with such contextual determination, giving rise on the one hand to Discourse Representation Theory or DRT (see Kamp, 1981, and the closely related File Change Semantics of Heim, 1982), and on the other to Situation Semantics (Barwise & Perry, 1983). Sperber and Wilson (working quite independently, judging from the absence of reference to all these developments) have pushed the issue somewhat further though: their claim is that not only is there wide pragmatic input into semantic interpretation (now generally acknowledged) but also that the VERY SAME PRAGMATIC PRINCIPLES (according to Sperber and Wilson, R of course) are involved in propositional determination and classic implicature generation (a view more likely to be resisted, for instance in Situation Semantics).[24] Sperber and Wilson would seem to be correct, although it is actually easier to show the systematicity of the intrusion of implicatures into semantic interpretation by looking at GCIs rather than Relevance implicatures (Levinson & Sag, in prep; Levinson, 1988).

If Sperber and Wilson are correct, the overall structure of the theory of meaning needs rethinking. Sperber and Wilson are Fodoreans, so they are forced into assuming that language is an input system delivering fragmentary, uninterpreted, logical forms to the central processor which is unable to feed information back to the input system. It follows that semantics, construed as part of the input system, has nothing to say about semantic interpretation, which actually belongs (like pragmatics and all nonspecialized processes) to the central processor. So virtually all the pre-occupations of modern

[24] The idea that implicature might play a crucial role in the determination of the proposition expressed by an utterance (and thus in the circumscription of its truth-conditions) is not in fact new. Indeed, L. Jonathan Cohen (1971) in an early critique of Grice, pointed out that this might be an unavoidable consequence of the Gricean programme. But the real credit for maintaining this view consistently when it was least popular surely belongs to Jay Atlas (1977; 1979: 275–279; 1989: 146 ff.). The idea was also gingerly explored by Wilson (1975: 150 ff.), and more robustly by Gazdar in the conclusions to his 1979 book; and, more recently and partly drawing on Sperber and Wilson's work, by Kempson (1986), Carston (1988), Levinson (1988), Levinson & Sag (in prep).

theoretical semantics will lie outside semantics proper; indeed they will not be distinguished from general thought processes, let alone from pragmatics. This conclusion is not necessary, and it is not natural. Instead, the evidence on which it is based constitutes good *prima facie* evidence against Fodoreanism, as does the previously mentioned difficulty with using R (a central process) to disambiguate the input as it comes in, so guiding the processing of the input device. We can maintain a systematic theory of meaning by adopting, as in garden-variety DRT, a level of representation to which both semantic and pragmatic modules can contribute, the whole being subject to semantic interpretation. This is theoretical modularity without Fodorean modularity. This solution is explored in largely independent work by Kadmon (1987), Levinson (1988) and van der Sandt (1988).

There are thus reasonably well worked out alternatives to the theory of R: theories that attempt to capture some of the content of pre-theoretical judgments of relevance, theories of non-monotonic inference, theories of generalized implicature that pre-empt Sperber and Wilson's 'explicatures', theories about the interaction of pragmatics and semantics in semantic interpretation. Each of these is closely attuned to intuition and data. True, these are not all bundled up as a single package. But then, for many of us, the discovery of lots of different aspects of meaning, each with its different properties, is exactly what has characterized recent progress in the theory of meaning.

4. Conclusions

This book is an ambitious bid for a paradigm-change in pragmatics. In the old paradigm, pragmatics is an untidy collection of usage principles, accrued over decades of careful observation, which together give some substantial account of uncoded utterance-meaning (see Levinson, 1983; Horn, 1988, for a survey). It may be a bit ramshackle, but it delivers the goods (or at least some of them); and new developments (as mentioned in the prior section) help to remedy deficiencies. In contrast, in the new paradigm we are being offered, pragmatics is reduced to a single cognitive principle, a mental reflex, which governs much else besides language use. Indeed, on the new paradigm, pragmatics (along with most of what is now considered semantics) disappears in a simple theory of general thought process (see Wilson & Sperber, 1986). Despite reassurances otherwise, it is obvious that phenomena central to one paradigm will be peripheral to the other; the coverage cannot be the same, indeed in the new paradigm it is not yet really clear what constitutes the fundamental data, for it is not a data-driven theory. Thus, the new paradigm offered here exists largely as manifesto. Some, like Isaiah Berlin's hedgehog, will find the bold simplicity attractive and will not be unduly worried by the gap between manifesto and programme; others, more like his fox, will be

infinitely suspicious, believing that nothing is simple in the theory of meaning, and they will find plenty here to be suspicious about.[25] Hedgehogs can diet on this book; foxes will be much cheered to hear that there is a marvellous feast of pragmatic complexities about to appear as antidote in Horn (1989). As Abraham Lincoln said of another book that galvanized opinion, 'People who like this sort of thing will find this the sort of thing they like.'

Author's address: Department of Linguistics,
University of Cambridge,
Sidgwick Avenue,
Cambridge CB3 9DA.

REFERENCES

Allen, J. (1983). Recognizing intentions from natural language utterances. In Brady, M. & Berwick, R. (eds) *Computational models of discourse*. Cambridge, Mass.: MIT Press. 107–164.
Atlas, J. (1977). Negation, ambiguity and presupposition. *Linguistics and Philosophy* 1. 321–336.
Atlas, J. (1979). How linguistics matters to philosophy: presupposition, truth and meaning. In Oh, C.-K. & Dinneen, D. (eds) *Syntax and semantics 11: Presupposition*. New York: Academic Press. 265–281.
Atlas, J. (1989). *Philosophy without ambiguity*. Oxford University Press.
Atlas, J. & Levinson, S. (1981). *It*-clefts, informativeness and logical form. In Cole, P. (ed.) *Radical pragmatics*. New York: Academic Press. 1–61.
Barwise, J. & Perry, J. (1983). *Situations and attitudes*. Cambridge, Mass.: MIT Press.
Berlin, I. (1978). *Russian thinkers*. London: The Hogarth Press.
Behavioral and Brain Sciences (BBS) 1987, **10**: 710–754. Sperber & Wilson: *Relevance*: Open Peer Commentary.
Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Blackwells.
Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
Carston, R. (1988). Implicature, explicature and truth-theoretic semantics. In Kempson, R. (ed.) 1988. *Mental representations: the interface between language and reality*. Cambridge: Cambridge University Press. 155–181.
Cohen, L. J. (1971). The logical particles of natural language. In Bar-Hillel, Y. (ed.) *Pragmatics of natural language*. Dordrecht: Reidel. 50–68.
Dascal, M. (1977). Conversational relevance. *Journal of Pragmatics* 1. 309–328.
Fodor, J. (1983). *Modularity of mind*. Cambridge, Mass.: MIT Press.
Gazdar, G. (1979). *Pragmatics: implicature, presupposition and logical form*. New York: Academic Press.
Gazdar, G. & Good, D. (1982). On a notion of relevance. In Smith, N. (ed.) *Mutual knowledge*. New York: Academic Press. 88–100.
Ginsberg, M. L. (ed.) (1987). *Readings in non-monotonic reasoning*. Los Altos, Ca.: Morgan Kaufmann.
Gumperz, G. (1983). *Discourse strategies*. Cambridge: Cambridge University Press.
Haegeman, L. (this volume). *Be going to* and *will*: a pragmatic account. *JL* **25**. 291–317.
Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Unpublished PhD. dissertation, University of Massachusetts, Amherst.
Hirschberg, J. (1985). *A theory of scalar implicature*. University of Pennsylvania, Moore School of Engineering, MS-CIS-85-56.

[25] 'There is a line among the fragments of the Greek poet Archilochus which says "The fox knows many things, but the hedgehog knows one big thing"... taken figuratively, the words can be made to yield a sense in which they mark one of the deepest differences which divide writers and thinkers... For there exists a great chasm between, on the one side, those who relate everything to a single central vision... and, on the other, those who pursue many ends' (Berlin, 1978: 22).

Hirschberg, J. (in press). *A theory of scalar implicature*. Cambridge: Cambridge University Press.
Hobbs, J. (1987). Implicature and definite reference. Stanford, CSLI Report NO. CSLI-87-99.
Hobbs, J. & Martin, P. (1987). Local pragmatics. *Proceedings, International Joint Conference on Artificial Intelligence*. Milan, August 1987. 520–523.
Holdcroft, D. (1987). Conversational relevance. In Verschueren & Bertuccelli-Papi (1987). 477–496.
Horn, L. (1972). On the semantic properties of the logical operators in English. Indiana University Linguistics Club.
Horn, L. (1983). Greek Grice. *Papers from the 9th Regional Meeting of the Chicago Linguistics Society*, 205–214.
Horn, L. (1984). Towards a new taxonomy for pragmatic inference. In Schiffrin, D. (ed.) *Meaning, form, and use in context*. Washington, D.C.: Georgetown University Press. 11–42.
Horn, L. (1985). Metalinguistic negation and pragmatic ambiguity. *Lg* **61**. 121–174.
Horn, L. (1988). Pragmatic theory. In Newmeyer, F. (ed.) *Linguistics: the Cambridge survey*, Vol. 1. Cambridge: Cambridge University Press. 113–145.
Horn, L. (1989). *A natural history of negation*. Chicago: Chicago University Press.
Johnson-Laird, P. (1983). *Mental models*. Cambridge: Cambridge University Press.
Kadmon, N. (1987). *On unique and non-unique reference and asymmetric quantification*. Unpublished PhD., University of Massachusetts, Amherst.
Kamp, H. (1981). A theory of truth and semantic representation. Reprinted in Groenendijk, J. & Janssen, T. & Stokhof, M. (1984). *Truth, interpretation and information*. Dordrecht: Foris. 1–41.
Kempson, R. (1986). Ambiguity and the semantics–pragmatics distinction, In Travis, C. (ed.) *Meaning and interpretation*, Oxford: Blackwells. 77–104.
Kempson, R. (1988). Grammar and conversational principles. In Newmeyer, F. (ed.) *Linguistics: the Cambridge survey*, Vol. 2. Cambridge: Cambridge University Press. 139–163.
Kempson, R. & Cormack, A. (1981). Ambiguity and quantification. *Linguistics & Philosophy* **4**. 259–309.
Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
Levinson, S. (1985). What's special about conversational inference? Unpublished paper presented to the Annual Meeting of the British Psychological Society, April 1985, University College, Swansea.
Levinson, S. (1987a). Pragmatics and the grammar of anaphora. *JL* **23**. 379–434.
Levinson, S. (1987b). Minimization and conversational inference. In Verschueren & Bertuccelli-Papi (1987). 61–130.
Levinson, S. (1988). Generalized conversational implicature and the semantics/pragmatics interface. MS. University of Cambridge.
Levinson, S. & Sag, I. (in prep.) Pragmatic inference and semantic interpretation.
Lewis, D. (1969). *Convention*. Cambridge, Mass.: Harvard University Press.
McCafferty, A. (1987). *Reasoning about implicature*. Unpublished PhD dissertation, University of Pittsburgh.
Pereira, F. & Pollack, M. (1988). An integrated framework for semantic and pragmatic interpretation. Menlo park: SRI International, Technical Note.
Perrault, C. R. (1987). An application of default logic to speech act theory. To appear in Cohen, P. Morgan, J. & Pollack, M. (eds) *Plans and intentions in communication and discourse*. Cambridge, Mass.: MIT Press.
Pollack, M. (1986). *Inferring domain plans in question-answering*. Menlo Park: SRI International, Technical Note 403.
Sadock, J. (1986). Remarks on the paper by Wilson and Sperber. *Papers from the parasession on pragmatics and grammatical theory. Chicago Linguistic Society*, 22, Part 2. 85–90.
Schiffer, S. (1972). *Meaning*. Oxford: Clarendon Press.
Seuren, P. (1988). The self-styling of Relevance Theory. *Journal of Semantics* **5**. 123–143.
Sidner, C. (1986). Focusing in the comprehension of definite anaphora. In Grosz, B., Sparck Jones, K., Webber, B. (eds) *Readings in natural language processing*. Los Altos, Ca.: Morgan Kaufmann. 363–394.
Silverstein, M. (1988). The indeterminacy of contextualization: when is enough enough? Paper presented to the conference 'The contextualization of language', October 1988, Konstanz.
Smith, N. (1981). Grammaticality, time and tense. In *The Psychological mechanisms of language*. London: The Royal Society and the British Academy. 39–51.

Smith, N. (ed.) (1982). *Mutual knowledge*. London: Academic Press.
Smith, N. & Smith, A. (1988). A relevance-theoretic account of conditionals. In Hyman, L. & Li, C. N. (eds) *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. London and New York: Routledge. 322–352.
Sperber, D. & Wilson, D. (1982). Mutual knowledge and relevance in theories of comprehension. In Smith, N. (ed.) (1982). 61–131.
Strawson, P. (1952). *Introduction to logical theory*. London: Methuen.
Thomason, R. (1987). Accommodation, meaning and implicatures: interdisciplinary foundations for pragmatics. MS, University of Pittsburgh. To appear in Cohen, P., Morgan, J., & Pollack, M. (eds) *Plans and intentions in communication and discourse*. Cambridge, Mass.: MIT Press.
Verschueren, J. & Bertuccelli-Papi, M. (eds) (1987). *The pragmatic perspective*. Amsterdam: John Benjamins.
van der Sandt, R. (1988). Discourse systems and echo-quotation. MS, University of Nijmegen, Dept. of Philosophy.
Wilson, D. (1975). *Presuppositions and non-truth-conditional semantics*. London: Academic Press.
Wilson, D. & Sperber, D. (1981). On Grice's theory of conversation. In Werth, P. *Conversation and discourse*. London: Croom Helm.
Wilson, D. & Sperber, D. (1986). Pragmatics and modularity. *Papers from the parasession on pragmatics and grammatical theory. Chicago Linguistics Society*, 22, Part 2. 67–84.
Zipf, G. K. (1949). *Human behavior and the principle of least human effort*. Cambridge, Mass.: Addison-Wesley.