

# 27 Relevance Theory

---

DEIRDRE WILSON AND DAN SPERBER

## 1 Introduction

Relevance theory may be seen as an attempt to work out in detail one of Grice's central claims: that an essential feature of most human communication is the expression and recognition of intentions (Grice 1989: Essays 1–7, 14, 18; Retrospective Epilogue). In elaborating this claim, Grice laid the foundations for an inferential model of communication, an alternative to the classical code model. According to the code model, a communicator encodes her intended message into a signal, which is decoded by the audience using an identical copy of the code. According to the inferential model, a communicator provides evidence of her intention to convey a certain meaning, which is inferred by the audience on the basis of the evidence provided. An utterance is, of course, a linguistically coded piece of evidence, so that verbal comprehension involves an element of decoding. However, the decoded linguistic meaning is just one of the inputs to a non-demonstrative inference process which yields an interpretation of the speaker's meaning.

The goal of inferential pragmatics is to explain how the hearer infers the speaker's meaning on the basis of the evidence provided. The relevance-theoretic account is based on another of Grice's central claims: that utterances automatically create expectations which guide the hearer toward the speaker's meaning. Grice described these expectations in terms of a Cooperative Principle and maxims of Quality (truthfulness), Quantity (informativeness), Relation (relevance), and Manner (clarity), which speakers are expected to observe (Grice 1961, 1989: 368–72). We share Grice's intuition that utterances raise expectations of relevance, but question several other aspects of his account, including the need for a Cooperative Principle and maxims, the focus on pragmatic contributions to implicit (as opposed to explicit) content, the role of maxim violation in utterance interpretation, and the treatment of figurative utterances.<sup>1</sup> The central claim of relevance theory is that the expectations of relevance raised by an utterance are precise and predictable enough to guide the hearer toward the speaker's meaning.

The aim is to explain in cognitively realistic terms what these expectations amount to, and how they might contribute to an empirically plausible account of comprehension. The theory has developed in several stages. A detailed version was published in *Relevance: Communication and Cognition* (Sperber and Wilson 1986a, 1987a, b) and updated in Sperber and Wilson (1995, 1998a, 2002) and Wilson and Sperber (2002). Here, we will outline the main assumptions of the current version of the theory and discuss some of its implications.

## 2 Relevance and Cognition

What sort of things may be relevant? Intuitively, relevance is a potential property not only of utterances and other observable phenomena, but of thoughts, memories, and conclusions of inferences. According to relevance theory, any external stimulus or internal representation which provides an input to cognitive processes may be relevant to an individual at some time. Utterances raise expectations of relevance not because speakers are expected to obey a Cooperative Principle and maxims or some other communicative convention, but because the search for relevance is a basic feature of human cognition, which communicators may exploit. In this section, we will introduce the basic notion of relevance and the Cognitive Principle of Relevance, which lay the foundation for the relevance-theoretic approach.

When is an input relevant? Intuitively, an input (a sight, a sound, an utterance, a memory) is relevant to an individual when it connects with background information he has available to yield conclusions that matter to him: say, by answering a question he had in mind, improving his knowledge on a certain topic, settling a doubt, confirming a suspicion, or correcting a mistaken impression. According to relevance theory, an input is relevant to an individual when its processing in a context of available assumptions yields a **POSITIVE COGNITIVE EFFECT**. A positive cognitive effect is a worthwhile difference to the individual's representation of the world: a true conclusion, for example. False conclusions are not worth having; they are cognitive effects, but not positive ones (Sperber and Wilson 1995: §3.1–2).

The most important type of cognitive effect is a **CONTEXTUAL IMPLICATION**, a conclusion deducible from input and context together, but from neither input nor context alone. For example, on seeing my train arriving, I might look at my watch, access my knowledge of the train timetable, and derive the contextual implication that my train is late (which may itself achieve relevance by combining with further contextual assumptions to yield further implications). Other types of cognitive effect include the strengthening, revision, or abandonment of available assumptions. For example, the sight of my train arriving late might confirm my impression that the service is deteriorating, or make me alter my plans to do some shopping on the way to work. According to relevance theory, an input is **RELEVANT** to an individual when, and only when, its processing yields such positive cognitive effects.<sup>2</sup>

Relevance is not just an all-or-none matter but a matter of degree. There are potentially relevant inputs all around us, but we cannot attend to them all. What makes an input worth picking out from the mass of competing stimuli is not just that it is relevant, but that it is MORE relevant than any alternative input available to us at that time. Intuitively, other things being equal, the more worthwhile conclusions achieved by processing an input, the more relevant it will be. According to relevance theory, other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater its relevance will be. Thus, the sight of my train arriving one minute late may make little worthwhile difference to my representation of the world, while the sight of it arriving half an hour late may lead to a radical reorganization of my day, and the relevance of the two inputs will vary accordingly.

What makes an input worth attending to is not just the cognitive effects it achieves. In different circumstances, the same stimulus may be more or less salient, the same contextual assumptions more or less accessible, and the same cognitive effects easier or harder to derive. Intuitively, the greater the effort of perception, memory, and inference required, the less rewarding the input will be to process, and hence the less deserving of attention. According to relevance theory, other things being equal, the greater the PROCESSING EFFORT required, the less relevant the input will be. Thus, RELEVANCE may be assessed in terms of cognitive effects and processing effort:

(1) **Relevance of an input to an individual**

- a. Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.
- b. Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time.

Here is a brief and artificial illustration of how the relevance of alternative inputs might be compared. Mary, who dislikes most meat and is allergic to chicken, rings her host to find out what is on the menu. He could truly tell her any of three things:

- (2) We are serving meat.
- (3) We are serving chicken.
- (4) Either we are serving chicken or  $(7^2 - 3)$  is not 46.

According to the characterization in (1), all three utterances would be relevant to Mary, but (3) would be more relevant than either (2) or (4). It would be more relevant than (2) for reasons of cognitive effect: (3) entails (2), and therefore yields all the conclusions derivable from (2), and more besides. It would be more relevant than (4) for reasons of processing effort: although (3) and (4) are logically equivalent, and therefore yield exactly the same cognitive effects,

these effects are easier to derive from (3) than from (4), which requires an additional effort of parsing and inference (in order to work out that the second disjunct is false and the first is therefore true). More generally, when similar amounts of effort are required, the effect factor is decisive, and when similar amounts of effect are achievable, the effort factor is decisive.

This characterization of relevance is comparative rather than quantitative: it allows clear comparisons in some cases, but not in all. While quantitative notions of relevance might be interesting from a formal point of view,<sup>3</sup> the comparative notion provides a better starting point for constructing a psychologically plausible theory. In the first place, only some aspects of effect and effort (e.g. processing time, number of contextual implications) are likely to be measurable in absolute numerical terms, while others (e.g. strength of implications, level of attention) are not. In the second place, even when absolute measures exist (for weight or distance, for example), we generally have access to more intuitive methods of assessment which are comparative rather than quantitative, and which are in some sense more basic. In therefore seems preferable to treat effort and effect (and relevance, which is a function of effort and effect) as **non-representational** dimensions of mental processes: they exist and play a role in cognition whether or not they are mentally represented; and when they are mentally represented, it is in the form of intuitive comparative judgments rather than absolute numerical ones.<sup>4</sup>

Within this framework, aiming to maximize the relevance of the inputs one processes is simply a matter of making the most efficient use of the available processing resources. No doubt this is something we would all want to do, given a choice. Relevance theory claims that humans do have an automatic tendency to maximize relevance, not because we have a choice in the matter – we rarely do – but because of the way our cognitive systems have evolved. As a result of constant selection pressures toward increasing efficiency, the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way. This universal tendency is described in the First, or Cognitive, Principle of Relevance (Sperber and Wilson 1995: §3.1–2):

(5) **Cognitive Principle of Relevance**

Human cognition tends to be geared to the maximization of relevance.

It is against this cognitive background that inferential communication takes place.

### 3 Relevance and Communication

The universal cognitive tendency to maximize relevance makes it possible (to some extent) to predict and manipulate the mental states of others. Knowing

your tendency to pick out the most relevant inputs and process them so as to maximize their relevance, I may be able to produce a stimulus which is likely to attract your attention, activate an appropriate set of contextual assumptions and point you toward an intended conclusion. For example, I may leave my empty glass in your line of vision intending you to notice and conclude that I might like another drink. As Grice pointed out, this is not yet a case of inferential communication because, although I intended to affect your thoughts in a certain way, I gave you no evidence that I had this intention. When I quietly leave my glass in your line of vision, I am not engaging in inferential communication, but merely exploiting your natural cognitive tendency to maximize relevance.

Inferential communication – what relevance theory calls *OSTENSIVE-INFERENTIAL COMMUNICATION*, for reasons that will shortly become apparent – involves an extra layer of intention:

(6) **Ostensive-inferential communication**

a. **The informative intention:**

The intention to inform an audience of something.

b. **The communicative intention:**

The intention to inform the audience of one's informative intention.<sup>5</sup>

Understanding is achieved when the communicative intention is fulfilled – that is, when the audience recognizes the informative intention. (Whether the informative intention itself is fulfilled depends on how much the audience trusts the communicator.)

How does the communicator indicate to an audience that she is trying to communicate in this overt, intentional way? Instead of covertly leaving my glass in your line of vision, I might touch your arm and point to my empty glass, wave it at you, ostentatiously put it down in front of you, stare at it meaningfully, or say, "My glass is empty." More generally, ostensive-inferential communication involves the use of an *OSTENSIVE STIMULUS*, designed to attract an audience's attention and focus it on the communicator's meaning. According to relevance theory, use of an ostensive stimulus may create precise and predictable expectations of relevance not raised by other inputs. In this section, we will describe these expectations and show how they may help to identify the communicator's meaning.

The fact that ostensive stimuli create expectations of relevance follows from the Cognitive Principle of Relevance. An ostensive stimulus is designed to attract the audience's attention. Given the cognitive tendency to maximize relevance, an audience will only pay attention to an input that seems relevant enough. By producing an ostensive stimulus, the communicator therefore encourages her audience to presume that it is relevant enough to be worth processing. This need not be a case of Gricean cooperation. Even a self-interested, deceptive, or incompetent communicator manifestly intends her audience to assume that her stimulus is relevant enough to be worth processing – why else would he

pay attention?<sup>6</sup> This is the basis for the Second, or Communicative, Principle of Relevance:

(7) **Communicative Principle of Relevance**

Every ostensive stimulus conveys a presumption of its own optimal relevance.

Use of an ostensive stimulus, then, creates a PRESUMPTION OF OPTIMAL RELEVANCE. The notion of optimal relevance is meant to spell out what the audience of an act of ostensive communication is entitled to expect in terms of effort and effect:

(8) **Presumption of optimal relevance**

- a. The ostensive stimulus is relevant enough to be worth the audience's processing effort.
- b. It is the most relevant one compatible with communicator's abilities and preferences.

According to clause (a), the audience can expect the ostensive stimulus to be at least relevant enough to be worth processing. Given the argument of section 2 that a stimulus is only worth processing if it is more relevant than any alternative input available at the time, this is not a trivial claim. Indeed, in order to satisfy the presumption of relevance, the audience may have to draw a stronger conclusion than would otherwise have been warranted. For example, if you just happen to notice my empty glass, you may be entitled to conclude that I *might* like a drink. If I deliberately wave it at you, you would generally be justified in concluding that I *would* like a drink.

According to clause (b), the audience of an ostensive stimulus is entitled to even higher expectations. The communicator wants to be understood. It is therefore in her interest – within the limits of her own capabilities and preferences – to make her ostensive stimulus as easy as possible for the audience to understand, and to provide evidence not just for the cognitive effects she aims to achieve but for further cognitive effects, which, by holding the audience's attention, will help her achieve her goal. For instance, the communicator's goal might be to inform her audience that she has started writing her paper. The most effective way of achieving this goal might be to offer more specific information and say, "I've already written a third of the paper." In the circumstances, her audience could then reasonably take her to mean that she has only written a third of the paper, because if she had written more, she should have said so, given clause (b) of the presumption of optimal relevance.

Of course, communicators are not omniscient, and they cannot be expected to go against their own interests and preferences. There may be relevant information that they are unable or unwilling to provide, and ostensive stimuli that would convey their intentions more economically, but that they are unwilling to produce, or unable to think of at the time. All this is allowed for in clause (b) of the presumption of optimal relevance, which states that the ostensive

stimulus is the most relevant one that the communicator is WILLING AND ABLE to produce (Sperber and Wilson 1995: §3.3 and 266–78).

This approach explains some parallels between ostensive and non-ostensive behavior that the Gricean framework obscures. Suppose you ask me a question and I remain silent. My silence may or may not be an ostensive stimulus. When it is not, you will naturally take it as indicating that I am unable or unwilling to answer; if I am clearly willing, you can conclude that I am unable, and if I am clearly able, you can conclude that I am unwilling. Given the presumption of optimal relevance, an ostensive silence can be analyzed as merely involving an extra layer of intention, and hence as COMMUNICATING – or IMPLICATING – that the addressee is unable or unwilling to answer.<sup>7</sup> In Grice's framework, however, violation of the first Quantity maxim invariably implicates INABILITY – rather than UNWILLINGNESS – to provide the required information. Inability to make one's contribution "such as is required" is consistent with the Cooperative Principle as long as it results from a clash with the Quality maxims. Unwillingness to make one's contribution "such as is required" is a violation of the Cooperative Principle; and since conversational implicatures are recoverable only on the assumption that the Cooperative Principle is being observed, it is impossible in Grice's framework to implicate that one is unwilling to provide the required information.<sup>8</sup> While cooperation in Grice's sense is quite common, we have argued that it is not essential to communication or comprehension (see note 6).

This account of communication has practical implications for pragmatics. The overall task of inferring the speaker's meaning may be broken down into a variety of pragmatic subtasks. There may be ambiguities and referential ambivalences to resolve, ellipses to interpret, and other underdeterminacies of explicit content to deal with. There may be implicatures to identify, illocutionary indeterminacies to resolve, metaphors and ironies to interpret. All this requires an appropriate set of contextual assumptions, which the hearer must also supply. The Communicative Principle of Relevance and the presumption of optimal relevance suggest a practical procedure for performing these subtasks and constructing a hypothesis about the speaker's meaning. The hearer should take the decoded linguistic meaning; following a path of least effort, he should enrich it at the explicit level and complement it at the implicit level until the resulting interpretation meets his expectation of relevance:

(9) **Relevance-theoretic comprehension procedure**

- a. Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.
- b. Stop when your expectations of relevance are satisfied (or abandoned).

Given clause (b) of the presumption of optimal relevance, it is reasonable for the hearer to follow a path of least effort because the speaker is expected (within the limits of her abilities and preferences) to make her utterance as

easy as possible to understand. Since relevance varies inversely with effort, the very fact that an interpretation is easily accessible gives it an initial degree of plausibility (an advantage specific to ostensive communication). It is also reasonable for the hearer to stop at the first interpretation that satisfies his expectations of relevance, because there should never be more than one. A speaker who wants her utterance to be as easy as possible to understand should formulate it (within the limits of her abilities and preferences) so that the first interpretation to satisfy the hearer's expectation of relevance is the one she intended to convey. An utterance with two apparently satisfactory competing interpretations would cause the hearer the unnecessary extra effort of choosing between them, and the resulting interpretation (if there were one) would not satisfy clause (b) of the presumption of optimal relevance.<sup>9</sup>

Thus, when a hearer following the path of least effort arrives at an interpretation that satisfies his expectations of relevance, in the absence of contrary evidence, this is the most plausible hypothesis about the speaker's meaning. Since comprehension is a non-demonstrative inference process, this hypothesis may well be false; but it is the best a rational hearer can do (on the role of the relevance-theoretic comprehension procedure in a modular approach to pragmatics, see section 5).

## 4 Relevance and Comprehension

In many non-verbal cases (e.g. pointing to one's empty glass, failing to answer a question), use of an ostensive stimulus merely adds an extra layer of intention recognition to a basic layer of information that the audience might have picked up anyway. In other cases (e.g. inviting someone out to a drink by miming the act of drinking), the communicator's behavior provides no direct evidence for the intended conclusion, and it is only the presumption of relevance that encourages the audience to spend the effort required to discover the communicator's meaning. In either case, the range of meanings that can be non-verbally conveyed is necessarily limited to those the communicator can evoke in her audience by drawing attention to observable features of the environment (whether pre-existing or produced specifically for this purpose).

In verbal communication, by contrast, speakers can convey a very wide range of meanings even though there is no independently identifiable basic layer of information for the hearer to pick up. What makes this possible is that utterances encode logical forms (conceptual representations, however fragmentary or incomplete) which the speaker has manifestly chosen to provide as input to the inferential comprehension process. As a result, verbal communication can achieve a degree of explicitness not available in non-verbal communication (compare pointing in the direction of a table containing glasses, ashtrays, plates, etc., and saying, "My glass is empty").

Although the decoded logical form of an utterance is an important clue to the speaker's intentions, it is now increasingly recognized that even the explicit



content of an utterance may go well beyond what is linguistically encoded.<sup>10</sup> What is still open to debate is how these context-dependent aspects of explicit content are recovered. Grice invoked his Cooperative Principle and maxims mainly to explain the recovery of implicatures,<sup>11</sup> and many pragmatists have followed him on this. There has thus been a tendency, even in much of the recent pragmatic literature, to treat the “primary” processes involved in the recovery of explicit content as significantly different from – i.e. less inferential, or less directly dependent on speakers’ intentions or pragmatic principles than – the “secondary” processes involved in the recovery of implicatures.<sup>12</sup>

In relevance theory, the identification of explicit content is seen as equally inferential, and equally guided by the Communicative Principle of Relevance, as the recovery of implicatures. The relevance-theoretic comprehension procedure applies in the same way to resolving linguistic underdeterminacies at both explicit and implicit levels. The hearer’s goal is to construct a hypothesis about the speaker’s meaning that satisfies the presumption of relevance conveyed by the utterance. As noted above, this overall task can be broken down into a number of subtasks:

(10) **Subtasks in the overall comprehension process**

- a. Constructing an appropriate hypothesis about explicit content (EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes.
- b. Constructing an appropriate hypothesis about the intended contextual assumptions (IMPLICATED PREMISES).
- c. Constructing an appropriate hypothesis about the intended contextual implications (IMPLICATED CONCLUSIONS).

These subtasks should not be seen as sequentially ordered: the hearer does not FIRST decode the logical form, THEN construct an explicature and select an appropriate context, and THEN derive a range of implicated conclusions. Comprehension is an on-line process, and hypotheses about explicatures, implicated premises, and implicated conclusions are developed in parallel against a background of expectations which may be revised or elaborated as the utterance unfolds.<sup>13</sup> In particular, the hearer may bring to the comprehension process not only a general presumption of relevance, but more specific expectations about how the utterance is intended to be relevant (what cognitive effects it is intended to achieve), and these may contribute, via backwards inference, to the identification of explicatures and implicated premises. Thus, each subtask in (10a–c) involves a non-demonstrative inference process embedded within the overall process of constructing a hypothesis about the speaker’s meaning.

To illustrate, consider the exchange in (11):

- (11)a. *Peter*: Did John pay back the money he owed you?
- b. *Mary*: No. He forgot to go to the bank.

Here is a schematic outline of how Peter might use the relevance-theoretic comprehension procedure in interpreting Mary's utterance, "He forgot to go to the bank":

(12)

<p>(a) Mary has said to Peter, "He<sub>x</sub> forgot to go to the BANK<sub>1</sub>/BANK<sub>2</sub>." [He<sub>x</sub> = uninterpreted pronoun] [BANK<sub>1</sub> = financial institution] [BANK<sub>2</sub> = river bank]</p>	<p><i>Embedding of the decoded (incomplete) logical form of Mary's utterance into a description of Mary's ostensive behavior.</i></p>
<p>(b) Mary's utterance will be optimally relevant to Peter.</p>	<p><i>Expectation raised by recognition of Mary's ostensive behavior and acceptance of the presumption of relevance it conveys.</i></p>
<p>(c) Mary's utterance will achieve relevance by explaining why John has not repaid the money he owed her.</p>	<p><i>Expectation raised by (b), together with the fact that such an explanation would be most relevant to Peter at this point.</i></p>
<p>(d) Forgetting to go to the BANK<sub>1</sub> may make one unable to repay the money one owes.</p>	<p><i>First assumption to occur to Peter which, together with other appropriate premises, might satisfy expectation (c). Accepted as an implicit premise of Mary's utterance.</i></p>
<p>(e) John forgot to go to the BANK<sub>1</sub>.</p>	<p><i>First enrichment of the logical form of Mary's utterance to occur to Peter which might combine with (d) to lead to the satisfaction of (c). Accepted as an explicature of Mary's utterance.</i></p>
<p>(f) John was unable to repay Mary the money he owes because he forgot to go to the BANK<sub>1</sub>.</p>	<p><i>Inferred from (d) and (e), satisfying (c) and accepted as an implicit conclusion of Mary's utterance.</i></p>
<p>(g) John may repay Mary the money he owes when he next goes to the BANK<sub>1</sub>.</p>	<p><i>From (f) plus background knowledge. One of several possible weak implicatures of Mary's utterance which, together with (f), satisfy expectation (b).</i></p>

Peter assumes in (12b) that Mary's utterance, decoded as in (12a), is optimally relevant to him. Since what he wants to know at this point is why John did not repay the money he owed, he assumes in (c) that Mary's utterance will achieve relevance by answering this question. One of the encoded logical forms provides easy access to the contextual assumption in (d) (that forgetting to go to the BANK<sub>1</sub> may prevent one repaying money one owes). This could be used as an implicit premise in deriving the expected explanation of John's behavior, as long as the utterance is interpreted on the explicit side (via disambiguation and reference resolution) as conveying the information in (e) (that John forgot to go to the BANK<sub>1</sub>). By combining the implicit premise in (d) and the explicit premise in (e), Peter arrives at the implicit conclusion in (f), from which further, weaker implicatures, including (g) and others, follow. The resulting interpretation satisfies Peter's expectations of relevance. Thus, explicatures and IMPLICATURES (implicit premises and conclusions) are arrived at by a process of mutual parallel adjustment, with hypotheses about both being considered in order of accessibility.<sup>14</sup>

This schematic outline of the comprehension process is considerably oversimplified.<sup>15</sup> In particular, it omits a range of lexical-pragmatic processes involved in the construction of explicatures. Consider the word *bank* in (11b). Peter would probably take this to denote not just a banking institution but a specific type of banking institution: one that deals with private individuals, and in particular, with John. Unless it is narrowed in this way, the explicit content of Mary's utterance will not warrant the conclusion in (12f), which is needed to satisfy Peter's expectation of relevance. (It is hard to see how the fact that John had forgotten to go to the World Bank, say, might explain his failure to repay the money he owed.) Similarly, he would take the phrase *go to the bank* to mean not merely visiting the bank, but visiting it in order to get money, and to get money in the regular way (legally, rather than, say, by robbing the bank). Unless the explicit content is narrowed in this way, it will not warrant the conclusion in (12f), which is needed to satisfy Peter's expectation of relevance.

Such stereotypical narrowings have sometimes been analyzed as generalized conversational implicatures or default interpretations, derivable via default rules.<sup>16</sup> Despite the richness and subtlety of much of this literature, relevance theory takes a different approach, for two main reasons. First, as noted above, it treats lexical narrowing as a pragmatic enrichment process which contributes to explicatures rather than implicatures.<sup>17</sup> Like all enrichment processes, narrowing is driven by the search for relevance, which involves the derivation of cognitive effects, and in particular of contextual implications. By definition, a contextual implication must follow logically from the explicatures of the utterance and the context. Sometimes, as in (11b), the explicit content must be contextually enriched in order to warrant the expected conclusion. In any framework where implicated conclusions are seen as logically warranted by explicit content, there is good reason to treat lexical narrowing as falling on the explicit rather than the implicit side.<sup>18</sup>

Second, lexical narrowing is much more flexible and context-dependent than appeals to generalized implicature or default interpretations allow. Barsalou (1987, 1992) surveys a range of experimental evidence which shows that even “stereotypical” narrowings of terms such as *bird*, *animal*, *furniture*, *food*, etc. vary across situations, individuals, and times, and are strongly affected by discourse context and considerations of relevance. Barsalou sees his results as best explained by assuming that lexical items give access not to ready-made prototypes (assignable by default rules) but to a vast array of encyclopedic assumptions, with different subsets being selected ad hoc to determine the occasion-specific interpretation of a word. On this approach, *bank* in (11b) might be understood as conveying not the encoded concept  $BANK_1$  but the ad hoc concept  $BANK^*$ , with a more restricted encyclopedic entry and a narrower denotation.

According to Barsalou, the process of ad hoc concept construction is affected by a range of factors including context, accessibility of encyclopedic assumptions, and considerations of relevance. The relevance-theoretic comprehension procedure may be seen as a concrete hypothesis about how such a flexible, relevance-governed lexical interpretation process might go. The hearer treats the linguistically encoded concept (e.g.  $BANK_1$  in (11b)) as no more than a clue to the speaker’s meaning. Guided by expectations of relevance, and using contextual assumptions (e.g. (12d)) made accessible by the encyclopedic entry of the linguistically encoded concept, he starts deriving cognitive effects. When he has enough to satisfy his expectations of relevance, he stops. The results would be as in (12) above, except that the contextual assumption in (d), the explicature in (e), and the implicatures in (f) and (g) would contain not the encoded concept  $BANK_1$  but the ad hoc concept  $BANK^*$ , with a narrower denotation, which would warrant the derivation of the expected cognitive effects.

The effect of such a flexible interpretation process may be a loosening rather than a narrowing of the encoded meaning (resulting in a broader rather than a narrower denotation). Clear cases include generic uses of prominent brand name (e.g. *Hoover*, *Xerox*, *Kleenex*) and loose uses of well-defined terms such as *square*, *painless*, or *silent*; but the phenomenon is very widespread. Consider *bank* in (11b). Given current banking practice, the word is sometimes loosely used to denote a category containing not only banking institutions but also automatic cash dispensers. Indeed, in order to satisfy his expectations of relevance in (11b), Peter would probably have to take it this way (i.e. to mean, roughly, “bank-or-cash-dispenser”). (If John regularly gets his money from a cash dispenser, the claim that he forgot to go to the  $BANK_1$ , might be strictly speaking false, and in any case would not adequately explain his failure to repay Mary.) Thus, *bank* in (11b) might be understood as expressing not the encoded concept  $BANK_1$ , but an ad hoc concept  $BANK^{**}$ , with a broader denotation, which shares with  $BANK_1$  the salient encyclopedic attribute of being a place where one goes in order to access money from one’s account. The interpretation of a quite ordinary utterance such as (11b) might then involve both a loosening and a narrowing of the encoded meaning.

Loose uses of language present a problem for Grice. Strictly speaking, faces are not square, rooms are generally not silent, and to describe them as such would violate his maxim of truthfulness (“Do not say what you believe to be false”). However, these departures from truthfulness do not fall into any of the categories of maxim-violation recognized by Grice (1989: 30). They are not covert violations, like lies, designed to deceive the hearer into believing what was said. They are not like jokes and fictions, which suspend the maxim entirely. Given their intuitive similarities to metaphor and hyperbole, it might be tempting to analyze them as overt violations (floutings), designed to trigger the search for a related implicature (in this case, a hedged version of what was said). The problem is that these loose uses are not generally perceived as violating the maxim of truthfulness at all. While we can all recognize on reflection that they are not strictly and literally true, these departures from truthfulness pass undetected in the normal flow of discourse. Grice’s framework thus leaves them unexplained.<sup>19</sup>

Loose uses are not the only problem for a framework with a maxim of truthfulness. There are questions about how the maxim itself is to be understood, and a series of difficulties with the analysis of tropes as overt violations (cf. Wilson and Sperber 2002). Notice, too, that the intuitive similarities between loose talk, metaphor, and hyperbole cannot be captured as long as metaphor and hyperbole are seen as overtly violating the maxim of truthfulness, while loose uses are not. We have argued that the best solution is to abandon the maxim of truthfulness and treat whatever expectations of truthfulness arise in utterance interpretation as by-products of the more basic expectation of relevance. On this approach, loose talk, metaphor, and hyperbole are merely alternative routes to achieving optimal relevance. Whether an utterance is literally, loosely, or metaphorically understood will depend on the mutual adjustment of content, context, and cognitive effects in order to satisfy the overall expectation of relevance.<sup>20</sup>

To illustrate, consider the exchange in (13):

- (13)a. *Peter*: What do you think of Martin’s latest novel?  
 b. *Mary*: It puts me to sleep.

Grice would treat Mary’s utterance in (13b) as having three distinct interpretations: as a literal assertion, a hyperbole, or a metaphor. Of these, Peter should test the literal interpretation first, and consider a figurative interpretation only if the literal interpretation blatantly violates the maxim of truthfulness. Yet there is now a lot of experimental evidence suggesting that literal interpretations are not necessarily tested and rejected before figurative interpretations are considered;<sup>21</sup> indeed, in interpreting (13b), it would probably not even occur to Peter to wonder whether Mary literally fell asleep.

Our analysis takes these points into account. In the first place, there is no suggestion that the literal meaning must be tested first. As with *bank* in (11b), the encoded conceptual address is merely a point of access to an ordered array of encyclopedic assumptions from which the hearer is expected to select.

Whether the resulting interpretation is literal or loose will depend on which assumptions he selects. In processing (13b), Peter will be expecting to derive an answer to his question: that is, an evaluation of the book. In the circumstances, a highly salient assumption will be that a book that puts one to sleep is likely to be extremely boring and unengaging. Having used this assumption to derive an answer that satisfies his expectations of relevance, he should stop. Just as in interpreting *bank* in (11b), it does not occur to him to wonder whether John gets his money from a bank or a cash dispenser, so in interpreting (13b), it should not occur to him to wonder whether the book literally puts Mary to sleep, almost puts her to sleep, or merely bores her greatly. Thus, the mutual adjustment process for (13b) should yield an explicature containing the ad hoc concept PUT TO SLEEP\*, which denotes not only literal cases of putting to sleep, but other cases that share with it the encyclopedic attribute of being extremely boring and unengaging. Only if such a loose interpretation fails to satisfy his expectations of relevance would Peter be justified in exploring further contextual assumptions, and moving toward a more literal interpretation.<sup>22</sup>

Generally, the explicit content of loose uses, and particularly of metaphors, is indeterminate to some degree. (Compare the concept SQUARE, SQUARE\*, AND SQUARE\*\* conveyed, respectively, by the literal phrase *square geometric figure*, the loose *square face*, and the metaphorical *square mind*.) This relative indeterminacy of explicatures is linked to the relative strength of implicatures.

A proposition may be more or less strongly implicated. It is STRONGLY IMPLICATED (or is a STRONG IMPLICATURE) if its recovery is essential in order to arrive at an interpretation that satisfies the addressee's expectations of relevance. It is WEAKLY IMPLICATED if its recovery helps with the construction of such an interpretation, but is not itself essential because the utterance suggests a range of similar possible implicatures, any one of which would do (Sperber and Wilson 1986a: 1.10–12, 4.6). For instance, (11b) strongly implicates (12f), since without this implication (or an appropriately narrowed-and-loosened variant), (11b) is not a relevant reply to (11a). (11b) also encourages the audience to derive a further implicature along the lines of (12g) (that John may repay Mary when he next goes to the bank), but here the audience must take some responsibility for coming to this conclusion rather than, say, the conclusion that John WILL repay Mary when he next goes to the bank, or some other similar conclusion.

Typically, loose uses, and particularly metaphorical uses, convey an array of weak implicatures. Thus, "John has a square mind" weakly implicates that John is somewhat rigid in his thinking, does not easily change his mind, is a man of principle, and so on. None of these implicatures is individually required for the utterance to make sense, but without some such implicatures, it will make no sense at all. If the word *square* is understood as expressing the concept SQUARE\*\*, which combines with contextual information to yield these implications, then the concept SQUARE\*\* itself will exhibit some indeterminacy or fuzziness, and the utterance as a whole will exhibit a corresponding weakness of explicature. Loose uses and metaphors typically exhibit such fuzziness, for which relevance theory provides an original account.

The distinction between strong and weak implicatures sheds light on the variety of ways in which utterance achieve relevance. Some utterances (e.g. technical instructions) achieve relevance by conveying a few strong implicatures. Others achieve relevance by weakly suggesting a wide array of possible implications, each of which is a weak implicature. This is typical of poetic uses of language, and has been discussed in relevance theory under the heading of *POETIC EFFECT* (Sperber and Wilson 1986a: 4.6–9, Pilkington 2000; for the related notions of *STYLISTIC EFFECT* and *PRESUPPOSITIONAL EFFECT*, see Sperber and Wilson 1986a: 4.5–6).

In Grice's framework (and indeed in all rhetorical and pragmatic discussions of irony as a figure of speech before Sperber and Wilson 1981) the treatment of verbal irony closely parallels the treatment of metaphor and hyperbole. For Grice, irony, like metaphor and hyperbole, is an overt violation of the maxim of truthfulness, differing only in the kind of implicature it conveys. We have argued not only against Grice's analysis of irony, but against the more general assumption that metaphor, hyperbole, and irony should be given parallel treatments.

Grice's account of irony is a variant of the classical rhetorical account on which an ironical utterance is seen as literally saying one thing and figuratively meaning the opposite. There are well-known arguments against this account. It is descriptively inadequate because ironical understatements, quotations, and allusions do not communicate the opposite of what is literally said. It is theoretically inadequate because saying the opposite of what one means is patently irrational; and this account does not explain why irony is universal and appears to arise spontaneously, without being taught or learned (Sperber and Wilson 1981, 1998b, Wilson and Sperber 1992).

According to relevance theory, verbal irony involves no special machinery or procedures not already needed to account for a basic use of language, *INTERPRETIVE USE*, and a specific form of interpretive use, *ECHOIC USE*.<sup>23</sup> An utterance may be interpretively used to (meta)represent another utterance or thought that it resembles in content. The best-known type of interpretive use is in reported speech or thought. An utterance is echoic when it achieves most of its relevance by expressing the speaker's attitude to views she tacitly attributes to someone else. Thus, suppose that Peter and Mary are leaving a party and one of the following exchanges occurs:

(14) *Peter*: That was a fantastic party.

(15) *Mary*: a. [happily] Fantastic.  
 b. [puzzled] Fantastic?  
 c. [scornfully] Fantastic!

In (15a), Mary echoes Peter's utterance in order to indicate that she agrees with it; in (15b), she indicates that she is wondering about it; and in (15c) she

indicates that she disagrees with it. The resulting interpretations might be as in (16):

- (16)a. She believes I was right to say/think that the party was fantastic.
- b. She is wondering whether I was right to say/think that the party was fantastic.
- c. She believes I was wrong to say/think that the party was fantastic.

Here, the basic proposition expressed by the utterances in (15) (*the party was fantastic*) is embedded under an appropriate higher-order speech-act or propositional-attitude description indicating, on the one hand, that the basic proposition is being used to interpret views Mary attributes to someone else, and, on the other, Mary's attitude to these attributed views. To understand Mary, Peter has to recognize not only the basic proposition expressed but also the fact that it is being attributively used, and the attitude Mary intends to convey.

The attitudes conveyed by an echoic utterance may be very rich and varied. The speaker may indicate that she endorses or dissociates herself from the views she is echoing: that she is puzzled, angry, amused, intrigued, skeptical, etc., or any combination of these. We treat verbal irony as involving the expression of a tacitly dissociative attitude – wry, skeptical, bitter, or mocking – to an attributed utterance or thought. Consider Mary's utterance in (15c) above. This is clearly both ironical and echoic. We claim that it is ironical BECAUSE it is echoic: verbal irony consists in echoing a tacitly attributed thought or utterance with a tacitly dissociative attitude.<sup>24</sup>

This approach sheds light on some cases of irony not adequately handled by the classical or Gricean accounts. Consider Mary's utterance, "He forgot to go to the bank," in (11b) above. There are situations where this might well be ironically intended even though it is neither blatantly false nor used to convey the opposite of what was said. Suppose Peter and Mary both know that John has repeatedly failed to repay Mary, with a series of pitifully inadequate excuses. Then (11b) may be seen as an ironical echo in which Mary tacitly dissociates herself from the latest excuse in the series. Thus, all that is needed to make (11b) ironical is a scenario in which it can be understood as a mocking echo of an attributed utterance or thought.<sup>25</sup>

One implication of this analysis is that irony involves a higher order of metarepresentational ability than metaphor. As illustrated in (16) above, the hearer of an echoic utterance must recognize that the speaker is thinking, not directly about a state of affairs but about a thought or utterance that she attributes to someone else. This implication of our account is confirmed by experimental evidence showing that irony comprehension requires second-order metarepresentational abilities, while metaphor comprehension requires only first-order abilities.<sup>26</sup> This difference is unexplained on the classical or Gricean accounts.<sup>27</sup>

Metarepresentational abilities also play a role in the interpretation of illocutionary acts. Consider the exchange in (17):



- (17)a. *Peter*: Will you pay back the money by Tuesday?  
 b. *Mary*: I will pay it back by then.

Both (17a) and (17b) express the proposition that *Mary will pay back the money by Tuesday*. In the interrogative (17a), this proposition is expressed but not communicated (in the sense that Peter does not put it forward as true, or probably true):<sup>28</sup> it is not an explicature of Peter's utterance. Yet intuitively, (17a) is no less explicit an act of communication than (17b). According to relevance theory, what is explicitly communicated by (17a) is the HIGHER-ORDER EXPLICATURE in (18):

- (18) Peter is asking Mary whether she will pay back the money by Tuesday.

Like all explicatures, (18) is recovered by a mixture of decoding and inference based on a variety of linguistic and non-linguistic clues (e.g. word order, mood indicators, tone of voice, facial expression).<sup>29</sup> In (17b), by contrast, the explicatures might include both (19a), the BASIC EXPLICATURE, and higher-order explicatures such as (19b) and (19c):

- (19)a. Mary will pay back the money by Tuesday.  
 b. Mary is promising to pay back the money by Tuesday.  
 c. Mary believes she will pay back the money by Tuesday.

Thus, an utterance may convey several explicatures, each of which may contribute to relevance and warrant the derivation of implicatures.<sup>30</sup>

On this approach, verbal irony has more in common with illocutionary and attitudinal utterances than it does with metaphor or hyperbole. As illustrated in (16c), the recognition of irony, like the recognition of illocutionary acts, involves the construction of higher-order explicatures, and therefore requires a higher degree of metarepresentational ability than the recognition of the basic proposition expressed by an utterance, whether literal, loose, or metaphorical.

More generally, on both Gricean and relevance-theoretic accounts, the interpretation of EVERY utterance involves a high degree of metarepresentational capacity, since overt communication involves a complex, multi-levelled mental state attribution (see section 3 above). This raises the question of how pragmatic abilities are acquired, and how they fit into the overall architecture of the mind.

## 5 Relevance Theory and Mental Architecture

Grice's analysis of overt communication treats comprehension as a variety of MIND-READING, or THEORY OF MIND (the attribution of mental states to others in order to explain and predict behavior).<sup>31</sup> The link between mind-reading and communication is confirmed by a wealth of developmental and neuropsychological evidence.<sup>32</sup> However, mind-reading itself has been analyzed in

rather different ways. Philosophers often describe it as an exercise in reflective reasoning (a central thought process, in Fodor's terms), and many of Grice's remarks about pragmatics are consistent with this. His rational reconstruction of how conversational implicatures are derived is a straightforward exercise in general-purpose "belief-desire" psychology:

He said that P; he could not have done this unless he thought that Q; he knows (and knows that I know that he knows) that I will realise that it is necessary to suppose that Q; he has done nothing to stop me thinking that Q; so he intends me to think, or is at least willing for me to think, that Q. (Grice 1989: 30–1)

In our own early work, we also treated pragmatic interpretation as a central, inferential process, albeit a spontaneous, intuitive rather than a conscious, reflective one (Sperber and Wilson 1986a: Chap. 2, Wilson and Sperber 1986b). More recently, there has been a tendency in the cognitive sciences to move away from Fodor's sharp distinction between modular input processes and relatively undifferentiated central processes and toward an increasingly modular view of the mind.<sup>33</sup> In this section, we will consider how the relevance-theoretic approach might fit with more modular accounts of inference, and in particular of mind-reading.<sup>34</sup>

One advantage of a domain-specific module is that it can contain special-purpose inferential procedures ("fast and frugal heuristics," in the terms of Gigerenzer et al. 1999) attuned to particular features of its own domain. In modular accounts of mind-reading, for example, standard "belief-desire" psychology is replaced by special-purpose inferential procedures justified by regularities existing only in this domain. Examples include an Eye Direction Detector, which infers perceptual and attentional states from direction of gaze, and an Intentionality Detector, which interprets self-propelled motion in terms of goals and desires (Leslie 1994, Premack and Premack 1994, Baron-Cohen 1995). This raises the question of whether there might be domain-specific communicative regularities to which a special-purpose comprehension module might be attuned.

Most approaches to mind-reading, whether modular or non-modular, assume that there is no need for special-purpose inferential comprehension procedures, because the mental-state attributions required for comprehension will be automatically generated by more general mechanisms which apply across the whole domain of intentional action (cf. Bloom 2000, 2002). However, there are problems with the view that speakers' meanings can be inferred from utterances by the same procedures used to infer intentions from actions. In the first place, the range of actions an agent can reasonably intend to perform in a given situation is in practice quite limited, and regular intention attribution is greatly facilitated by this. By contrast, the range of meanings a speaker can reasonably intend to convey in a given situation is virtually unlimited (cf. section 3 above). It is simply not clear how the standard procedures for intention attribution could yield attributions of speakers' meanings except in easy and trivial cases (Sperber 2000, Sperber and Wilson 2002).

In the second place, inferential comprehension typically involves several layers of metarepresentation (cf. sections 4 and 5 above), while in regular mind-reading a single level is generally enough. This discrepancy is particularly apparent in child development. It is hard to believe that two-year-old children, who fail, for instance, on regular first-order false belief tasks, can recognize and understand the peculiar multi-levelled representations involved in overt communication, using nothing more than a general-purpose “belief–desire” psychology. This makes the possibility of a special-purpose comprehension submodule worth exploring (Sperber 1996, 2000, 2002, Origg and Sperber 2000, Wilson 2000, Sperber and Wilson 2002).

We have argued (following Sperber 1996) that the regularity described in the Communicative Principle of Relevance (that acts of ostensive communication create presumptions of relevance) underpins the workings of a special-purpose inferential comprehension device. On this approach, the relevance-theoretic comprehension procedure in (9) should be seen not as a variant of Grice’s working-out schema, but as a dedicated inferential mechanism, a “fast and frugal heuristic,” which automatically computes a hypothesis about the speaker’s meaning on the basis of the linguistic and other evidence provided.

This approach allows for varying degrees of sophistication in the hearer’s expectations of relevance. In an unsophisticated version (presumably the one always used by young children), what is expected is actual optimal relevance. In a more sophisticated version (used by competent adult communicators who are aware that the speaker may be mistaken about what is relevant to the hearer, or in bad faith and merely intending to seem relevant), what is expected may be merely attempted or purported optimal relevance. Adult communicators may nevertheless expect actual optimal relevance by default (Sperber 1994, Bezuidenhout and Sroda 1998, Wilson 2000, Happé and Loth 2002).

The complexity of the inferences required by Grice’s account of communication has sometimes been seen as an argument against the whole inferential approach. We are suggesting an alternative view on which, just as children do not have to learn their language but come with a substantial innate endowment, so they do not have to learn what ostensive-inferential communication is, but come with a substantial innate endowment.

## 6 Conclusion: An Experimentally Testable Cognitive Theory

Relevance theory is a cognitive psychological theory. Like other psychological theories, it has testable consequences: it can suggest experimental research, and is open to confirmation, disconfirmation, or fine-tuning in the light of experimental evidence. As with other theories of comparable scope, its most general claims can be tested only indirectly. For example, the Cognitive Principle of Relevance suggests testable predictions only when combined with descriptions

of particular cognitive mechanisms (e.g. for perception, categorization, memory, or inference). Given a description of such a mechanism, it may be possible to test the relevance-theoretic claim that this mechanism contributes to a greater allocation of cognitive resources to potentially relevant inputs, by comparing it with some alternative hypothesis, or at least the null hypothesis.

The Communicative Principle of Relevance is a law-like generalization which follows from the Cognitive Principle of Relevance, combined with a broadly inferential view of communication. It could be falsified by finding genuine communicative acts which do not convey a presumption of optimal relevance (but rather, say, a presumption of literal truthfulness, or maximal informativeness, or no such presumption at all). When combined with descriptions of specific types and properties of communicative acts, it yields precise predictions, some of which have been experimentally tested.

In this survey, we have drawn attention to several cases where the predictions of relevance theory differ from those more or less clearly suggested by alternative frameworks, and where the relevance-theoretic analyses have been experimentally tested and their predictions confirmed. We will end with two further illustrations of how his approach yields testable predictions.

As noted in section 2 above, relevance theory does not provide an absolute measure of mental effort or cognitive effect, and it does not assume that such a measure is available to the spontaneous workings of the mind. What it does assume is that the actual or expected relevance of two inputs can quite often be compared. These possibilities of comparison help individuals to allocate their cognitive resources, and communicators to predict and influence the cognitive processes of others. They also enable researchers to manipulate the effect and effort factors in experimental situations.

For example, consider the conditional statement in (20), describing a series of cards with letters or numbers on both front and back:

(20) If a card has a 6 on the front, it has an E on the back.

In the Wason selection task (the most famous experimental paradigm in the psychology of reasoning; cf. Wason 1966), participants are shown four cards with (say) a 6, a 4, an E, and an A on the front, and asked which ones they would have to turn over to check whether (20) is true or false. The correct response is to select the 6 and the A cards. By 1995, literally thousands of experiments with similar materials had failed to produce a majority of correct responses. Most people choose either the 6 card alone, or the 6 and the E. In "Relevance theory explains the selection task" (1995), Sperber, Cara, and Girotto argued that participants interpret conditional statement by deriving testable implications in order of accessibility, stop when their expectations of relevance are satisfied, and choose cards on the basis of this interpretation. Using this idea, Sperber et al. were able, by varying the content and context of (20), to manipulate the effort and effect factors so as to produce correct or incorrect selections at will.

Typically, a conditional statement of the form *If P then Q* achieves relevance by allowing the consequent *Q* to be derived whenever the antecedent *P* is satisfied. With (20), this leads to selection of the 6 card. Another common way for a conditional statement to achieve relevance is by creating the expectation that both *P* and *Q* are true. With (20), this leads to selection of the 6 and E cards. Of course, a conditional also implies that *P* and *not-Q* will not be true together. By choosing cards on this basis, participants would correctly select the 6 and A cards. However, in most contexts this implication is relatively costly to derive, leads to no further effects, and would not be derived by a hearer looking for optimal relevance. What Sperber et al. did was to manipulate the effort and effect factors (either separately or together), to make this implication easier and/or more rewarding to derive, and the correct cards correspondingly more likely to be chosen. In the most successful condition, (20) was presented as a statement made by an engineer who has just repaired a machine which was supposed to print cards according to the specification in (20), but which had malfunctioned and wrongly printed cards with a 6 on the front and an A on the back. Here, (20) achieved relevance by implying that there would be no more cards with a 6 on the front and an A rather than an E on the back, and a majority of participants made the correct selection. This experiment shows that performance on the selection task is determined not merely by general-purpose or special-purpose reasoning abilities (as had generally been assumed) but by pragmatic factors affecting the interpretation of conditional statements. It also confirms that the interpretation of conditionals is governed by the twin factors of effort and effect, either separately or in combination.<sup>35</sup>

Here is a second example of how the interaction of effort and effect can be experimentally investigated, this time in utterance production rather than interpretation. Suppose a stranger comes up and asks you the time. You look at your watch and see that it is 11:58 exactly. How should you reply? In Grice's framework, a speaker obeying the maxim of truthfulness should say "11:58." By saying "It's 12:00" (speaking loosely and thus violating the maxim of truthfulness), you would be understood as conveying that it was (exactly) 12:00. By contrast, a speaker aiming at optimal relevance has every reason to speak loosely (thus reducing her hearer's processing effort) unless this would (in her view) lead to some significant loss of cognitive effects. It should therefore be possible, by varying the scenario in which the question is asked, to produce stricter or looser answers depending on whether or not the stricter answer would carry relevant implications. This prediction has been experimentally tested, and the relevance-theoretic analysis confirmed: strangers in public places asked for the time tend to speak loosely or give strictly accurate answers depending on subtle clues as to what might make it relevant for the questioner to know the time (van der Henst et al. 2002).

So far, the main obstacle to experimental comparisons of relevance theory with other pragmatic theories has been that the testable consequences of these other theories have not always been explicitly spelled out. Most pragmatic research has been carried out in a philosophical or linguistic tradition, which

places a higher priority on theoretical generality and reliance on intuitions than on the need for experimentation. Relevance theorists have been trying to combine theoretical generality with all the possibilities of testing provided by the careful use of linguistic intuitions, observational data, and the experimental methods of cognitive psychology. We see this as an important direction for future research.

---

## ACKNOWLEDGMENTS

---

We are grateful to Larry Horn and Gregory Ward for valuable comments and suggestions, and to the many friends, colleagues, and students whose positive proposals and criticisms have contributed greatly to the development of the theory.

---

## NOTES

---

- 1 For early arguments against these aspects of Grice's framework, see Sperber and Wilson (1981), and Wilson and Sperber (1981).
- 2 For early accounts of COGNITIVE (or CONTEXTUAL) EFFECTS, see Wilson and Sperber (1981, 1986b). For the standard definitions, see Sperber and Wilson (1986a: 2.7, especially note 26). On the deductive inferences involved, see Politzer (1990) and Sperber and Wilson (1990a). There may be still further types of positive cognitive effect (improvements in memory or imagination, for example; cf. Wilson and Sperber 2002).
- 3 For suggestions about how this might be done, see Sperber and Wilson (1986a: 124–32). Formal notions of relevance have been explored by Merin (1999), Blutner (1998) (which brings together ideas from Horn 1984a, 1992, Levinson 1987a, 2000a, Hobbs et al. 1993, and Sperber and Wilson 1986a); van Rooy (1999, 2001).
- 4 On COMPARATIVE and QUANTITATIVE concepts, see Sperber and Wilson (1986a: 79–81, 124–32). On factors affecting comparative and quantitative assessments of relevance, see Sperber and Wilson (1986a): 3.2, 3.5, 3.6.
- 5 This is the simpler of two characterizations of ostensive-inferential communication in Sperber and Wilson (1986a: 29, 58, 61). The fuller version involves the notions of MANIFESTNESS and MUTUAL MANIFESTNESS. We argue that for communication to be truly overt, the communicator's informative intention must become not merely manifest to the audience (i.e. capable of being recognized and accepted as true, or probably true), but mutually manifest to communicator and audience. On the communicative and informative intentions, see Sperber and Wilson (1986a): 1.9–12; on mutual manifestness, see Garnham and Perner (1990) and Sperber and Wilson (1990a).

- 6 On Gricean cooperation and communication, see Kasher (1976), Wilson and Sperber (1981), Sperber and Wilson (1986a: 161–2), Smith and Wilson (1992), Sperber (1994, 2000), Sperber and Wilson (2002).
- 7 On ostensive silences, see Morgan and Green (1987: 727) and Sperber and Wilson (1987b: 746–7).
- 8 The symmetry between unwillingness and inability to provide relevant information is also lost in Gricean analyses of scalar implicatures. See Sperber and Wilson (1995: 276–8), Green (1995), Matsumoto (1995), Carston (1995, 1998b), and section 6 below. For experimental work, see Noveck (2001), Papafragou (2002), Papafragou and Musolino (2002).
- 9 Puns and deliberate equivocations are sometimes seen as creating problems for this approach (e.g. Morgan and Green 1987: 726–7). We would analyze them as cases of layering in communication. Just as failure to provide relevant information at one level may be used as an ostensive stimulus at another, so production of an utterance which is apparently uninterpretable at one level may be used as an ostensive stimulus at another (Sperber and Wilson 1987b: 751, Tanaka 1992).
- 10 By “explicitly communicated content” (or EXPLICATURE), we mean a proposition recovered by a combination of decoding and inference, which provides a premise for the derivation of contextual implications and other cognitive effects (Sperber and Wilson 1986a: 176–93, Carston 2002b; this volume). Despite many terminological disagreements (see notes 17 and 18), the existence of pragmatic contributions at this level is now widely recognized (Wilson and Sperber 1981, 1998, 2002, Kempson and Cormack 1982, Travis 1985, 2001, Sperber and Wilson 1986a: 4.2–3, Kempson 1986, 1996, Blakemore 1987, Carston 1988, 2000, 2002a, 2002b, Recanati 1989, 2002a, Neale 1992, Bach 1994a, 1994b, 1999a, Stainton 1994, 1997b, this volume, Bezuidenhout 1997, Levinson 2000a, Fodor 2001).
- 11 In his “Retrospective Epilogue,” and occasionally elsewhere, Grice seems to acknowledge the possibility of intentional pragmatic contributions to “dictive content” (Grice 1989: 359–68). See Carston (2002b) and Wharton (in preparation).
- 12 On primary and secondary pragmatic processes, see Breheny (2002), Recanati (2002b), Carston (2002b, this volume), and Sperber and Wilson (2002). Some work on generalized conversational implicature and discourse pragmatics tacitly invokes a similar distinction (cf. Hobbs 1985b, Lascarides and Asher 1993, Lascarides et al. 1996, Levinson 2000a). See also notes 17 and 18.
- 13 See Sperber and Wilson (1986a): 4.3–5, esp. pp. 204–8, and Wilson and Sperber (2002).
- 14 For ease of exposition, we have used an example where preceding discourse creates a specific expectation of relevance, so that the interpretation process is strongly driven by expectations of effect. In an indirect answer such as (ib), where there are two possible implicatures (positive or negative), considerations of effort, and in particular the accessibility of contextual assumptions, play a more important role. In a discourse-initial utterance such as (ii), or in a questionnaire situation, considerations of effort are likely

- to play a decisive role in choosing among possible interpretations:
- (i)a. *Peter*: Did John pay back the money he owed?
  - b. *Mary*: He forgot to go to the bank.
  - (ii) He forgot to go to the bank.
- 15 For one thing, the assumptions that Peter entertains in interpreting the utterance are presumably not represented in English but in some conceptual representation system or language of thought. We also ignore semantic issues such as the analysis of definite articles and definite descriptions (e.g. *the bank*).
- 16 See e.g. Horn (1984a, 1992), Levinson (1987a, 2000a), Hobbs et al. (1993), Lascarides et al. (1996), Lascarides and Copestake (1998), Blutner (1998, to appear).
- 17 As noted above (note 10), there is some debate about how the explicit–implicit distinction should be drawn (e.g. Horn 1992, Sperber and Wilson 1986a: 4.1–4, Wilson and Sperber 1993, Bach 1994a, 1994b, 1999a, Levinson 2000a, Carston 2002a, 2002b, this volume). The issue is partly terminological, but becomes substantive when combined with the claim that explicit and implicit communication involve distinct pragmatic processes (as in much of the literature on generalized implicatures, cf. Levinson 2000a).
- 18 Levinson (2000a: 195–6) rejects the explicature–implicature distinction on the ground that no criterion for distinguishing explicatures from implicatures is provided. Our notion of an explicature is motivated, among other things, by embedding tests which suggest that certain pragmatic processes contribute to truth-conditional content, while others do not (Wilson and Sperber 1986b: 80; 2002, Ifantidou 2001). The allocation of pragmatically inferred material between explicatures and implicatures is constrained, on the one hand, by our theoretical definitions of explicature and implicature (Sperber and Wilson 1986a: 182, Carston 2002b), and, on the other, by the fact that implicated conclusions must be warranted by the explicit content, together with the context. See Sperber and Wilson (1986a: 4.3), Sperber and Wilson (1998a), Carston (1995, 1998b, 2000, 2002b, this volume), Wilson and Sperber (1998, 2002). For experimental evidence, see Gibbs and Moise (1997), Matsui (1998, 2000), Nicolle and Clark (1999), Wilson and Matsui (2000), Noveck (2001), Papafragou (2002), Papafragou and Musolino (2003).
- 19 Over time, lexical loosening may stabilize in a community and give rise to an extra sense, which may in turn be narrowed or loosened for occasion-specific pragmatic reasons. Typically, there are too many occasion-specific interpretations to allow a purely semantic or default-pragmatic account (Searle 1979, 1980, Horn 1984a, G. Lakoff 1987, Franks and Braisby 1990, Sweetser 1990, Hobbs et al. 1993, Bach 1994a, 1994b, 1999a, Recanati 1995, Carston 1997, 1998b, 2002b, this volume, Sperber and Wilson 1998a, Wilson 1998, Lasersohn 1999, Traugott 1999, Asher and Lascarides 2001, Papafragou 2000, Wilson and Sperber 2002).
- 20 For early arguments against the maxim of truthfulness, see Wilson and Sperber (1981). For a detailed critique, see Wilson and Sperber (2002). For experimental evidence, see Matsui (1998, 2000), Wilson and Matsui (2000), van der Henst et al. (2002).



- 21 See e.g. Gibbs (1994), Glucksberg (2001), Noveck et al. (2001). Glucksberg's claim that metaphor interpretation involves constructing a more inclusive category fits well with our account.
- 22 While the claim that metaphor is a variety of loose use has been part of the theory for some time (e.g. Sperber and Wilson 1985/6, 1986a: 4.7–8, 1990b), some details of this analysis are more recent. For discussion, see Recanati (1995), Carston (1997, 2002b, this volume), Sperber and Wilson (1998a), Wilson and Sperber (2002).
- 23 On interpretive use, see Sperber and Wilson (1986a: 4.7), Blass (1990), Gutt (1991), Sperber (1997), Papafragou (1998, 2000), Wilson (2000), Noh (2001). On the echoic use, see Sperber and Wilson (1986a: 4.9), Blakemore (1994), Carston (1996, 2002b), Noh (1998), Wilson (2000).
- 24 This account of irony was first proposed in Sperber and Wilson 1981, and extended in Sperber and Wilson (1986a: 4.7, 4.9), Sperber and Wilson (1990b, 1998b), Wilson and Sperber (1992), Curcò (1998). For critical discussion, see Clark and Gerrig (1984), Kreuz and Glucksberg (1989), Gibbs and O'Brien (1991), Martin (1992), Kumon-Nakamura et al. (1995), and the papers by Seto, Hamamoto, and Yamanashi in Carston and Uchida (1998). For responses, see Sperber (1984), Sperber and Wilson (1998b).
- 25 For experimental evidence, see Jorgensen et al. (1984), Happé (1993), Kreuz and Glucksberg (1989), Gibbs and O'Brien 1991, Gibbs (1994), Kumon-Nakamura et al. (1995), Langdon et al. (2002).
- 26 On the development of metaphor and irony, see Winner (1988).
- On the relation between irony, metaphor, and metarepresentational abilities, see Happé (1993), Langdon et al. (2002). On communicative and metarepresentational, see section 5 below.
- 27 Levinson (2000a: 239), who interprets us (mistakenly) as claiming that irony does not contribute to explicatures, objects that we cannot account for the fact that ironical use of a referential expression may make a difference to truth conditions (as in his nice example "If you need a car, you may borrow my Porsche" [used to refer to the speaker's VW]). In fact, we treat irony as a variety of free indirect speech, which is closely related to metalinguistic use and contributes directly to explicatures. It is uncontroversial that free indirect speech and metalinguistic use may make a difference to truth conditions (Horn 1989, Sperber and Wilson 1981, 1986a: 4.7, Carston 1996, 2002b, Cappelen and Lepore 1997, Noh 2000, Wilson 2000), and Levinson's example fits well with our account.
- 28 See Sperber and Wilson (1986a): 1.9–12.
- 29 Mood indicators are among the items seen in relevance theory as carrying procedural rather than conceptual meaning. See Blakemore (1987, 2002, this volume), Wharton (2001, in press, in preparation), and Iten (2000b).
- 30 On higher-level explicatures, see Blakemore (1991), Wilson and Sperber (1993), and Ifantidou (2001). On non-declarative utterances, see Sperber and Wilson (1986a: 4.10), Wilson and Sperber (1988), Wilson (2000), and Noh (2001). For critical discussion, see Bird (1994) and Harnish (1994).

- 31 See Whiten (1991), Davies and Stone (1995a, 1995b), Carruthers and Smith (1996), and Malle et al. (2001).
- 32 See Perner et al. (1989), Happé (1993), Baron-Cohen (1995), Mitchell et al. (1999), Happé and Loth (2002), Papafragou (2002), and *Mind and Language* 17.1–2 (2002).
- 33 We use “module” in a somewhat broader sense than Fodor’s, to mean a domain-specific autonomous computational mechanism (cf. Sperber 1996: Chapter 6, 2002).
- 34 See Leslie (1991), Hirschfeld and Gelman (1994), Barkow et al. (1995), Sperber (1996, 2002), and Fodor (2000).
- 35 For other experiments on the selection task, see Girotto et al. (2001), Sperber and Girotto (to appear). For other applications of relevance theory to the psychology of reasoning, see van der Henst (1999), Politzer and Macchi (2000), and van der Henst et al. (2002).