Truthfulness and Relevance

DEIRDRE WILSON & DAN SPERBER

Abstract

This paper questions the widespread view that verbal communication is governed by a maxim, norm or convention of literal truthfulness. Pragmatic frameworks based on this view must explain the common occurrence and acceptability of metaphor, hyperbole and loose talk. We argue against existing explanations of these phenomena and provide an alternative account, based on the assumption that verbal communication is governed not by expectations of truthfulness but by expectations of relevance, which are raised by literal, loose and metaphorical talk alike. Sample analyses are provided, and some consequences of this alternative account are explored.

1 Introduction

Here are a couple of apparent platitudes. As speakers, we expect what we say to be accepted as true. As hearers, we expect what is said to us to be true. If it were not for these expectations, if they were not often enough satisfied, there would be little point in communicating at all. David Lewis (who has proposed a convention of truthfulness) and Paul Grice (who has argued for maxims of truthfulness), among others, have explored some of the consequences of these apparent platitudes. We want to take a different line and argue that they are strictly speaking false. Of course hearers expect to be informed and not misled by what is communicated; but what is communicated is not the same as what is said. We will argue that language use is not governed by any convention or maxim of truthfulness. Whatever genuine facts such a convention or maxim was supposed to explain are better explained by assuming that communication is governed by a principle of relevance.

According to David Lewis (1975; 1983), there is a regularity (and a moral obligation) of truthfulness in linguistic behaviour. This is not a convention in Lewis's sense, since there is no alternative regularity which would be preferable as long as everyone conformed to it. However, for any language £ of a population P, Lewis argues that there is a convention of *truthfulness and trust in £* (an alternative being a convention of truthfulness and trust in some other language £'):

> My proposal is that the convention whereby a population P uses a language £ is a convention of truthfulness and trust in £. To be truthful in £ is to act in a certain way: to try never to utter any sentences of £ that are not true in £. Thus it is to avoid uttering any sentence of £ unless one believes it to be true in £. To be trusting in £ is to form beliefs in a certain way: to impute truthfulness in £ to others, and thus to tend to respond to another's utterance of any sentence of £ by coming to believe that the uttered sentence is true in £. (Lewis 1983: 167)

Lewis considers the objection that truthfulness might not be the only factor that needs to be taken into account, and replies as follows:

> *Objection*: Communication cannot be explained by conventions of truthfulness alone. If I utter a sentence σ of our language £, you—expecting me to be truthful in £—will conclude that I take σ to be true in £. If you think I am well informed, you will also conclude that probably σ is true in £. But you will draw other conclusions as well, based on your legitimate assumption that it is for some good reason that I chose to utter σ rather than remain silent, and rather than utter any of the other sentences of £ that I also take to be true in £. I can communicate all sorts of misinformation by exploiting your beliefs about my conversational purposes, without ever being untruthful in £. Communication depends on principles of helpfulness and relevance as well as truthfulness.
> *Reply*: All this does not conflict with anything I have said. We do conform to conversational regularities of helpfulness and relevance. But these regularities are not independent conventions of language; they result from our convention of truthfulness and trust in £ together with certain general facts—not dependent on any convention—about our conversational purposes and our beliefs about one another. Since they are by-products of a convention of truthfulness and trust, it is unnecessary to mention them separately in specifying the conditions under which a language is used by a population.
> (Lewis 1983: 185)

However, Lewis does not explain how regularities of relevance might be by-products of a convention of truthfulness. One of our aims will be to show that, on the contrary, expectations of truthfulness are a by-product of expectations of relevance.

Paul Grice, in his *William James Lectures*, sketched a theory of utterance interpretation based on a Co-operative Principle and maxims of truthfulness, informativeness, relevance and clarity (Quality, Quantity, Relation and Manner). The Quality maxims went as follows:

(1) *Grice's maxims of Quality*
Supermaxim: Try to make your contribution one that is true.
(i) Do not say what you believe to be false. [*maxim of truthfulness*]
(ii) Do not say that for which you lack adequate evidence.

The supermaxim of Quality is concerned with the speaker's overall contribution (what is communicated, either explicitly or implicitly), while the first and second maxims of Quality relate only to what is said (the proposition explicitly expressed or asserted). Grice saw the first maxim of Quality, which we will call the maxim of truthfulness, as the most important of all the maxims. He says in the *William James Lectures*:

> It is obvious that the observance of some of these maxims is a matter of less urgency than is the observance of others; a man who has expressed himself with undue prolixity would, in general, be open to milder comment than would a man who has said something he believes to be false. Indeed, it might be felt that the importance of at least the first maxim of Quality is such that it should not be included in a scheme of the kind I am constructing; other maxims come into operation only on the assumption that this maxim of Quality is satisfied. While this may be correct, so far as the generation of implicatures is concerned, it seems to play a role not totally different from the other maxims, and it will be convenient, for the present at least, to treat it as a member of the list of maxims. (Grice 1989: 27)

In his "Retrospective Epilogue", written 20 years later, this view is apparently maintained:

> The maxim of Quality, enjoining the provision of contributions which are genuine rather than spurious (truthful rather than mendacious), does not seem to be just one among a number of recipes for producing contributions; it seems rather to spell out the difference between something's being and (strictly speaking) failing to be, any kind of contribution at all. False information is not an inferior kind of information; it just is not information.
> (Grice 1989: 371)

Notice, though, an interesting shift. While he talks of "the maxim of Quality," Grice's concern here is with the speaker's contribution as a whole; indeed, there is room for doubt about whether he had the first maxim of Quality or the supermaxim in mind. We believe that this is not a minor detail. One of our aims is to show that the function Grice attributes to the Quality maxims—ensuring the quality of the speaker's overall contribution—can be more effectively achieved in a framework with no maxim of truthfulness at all.

There is a range of apparent counterexamples to the claim that speakers try to tell the truth. These include lies, jokes, fictions, metaphors and ironies. Lewis and Grice are well aware of these cases, and discuss them in some detail. Grice (1989: 30), for instance, notes that his maxims may be violated, and lists several categories of violation, each with its characteristic effects. Lies are examples of *covert violation*, where the hearer is meant to assume that the maxim of truthfulness is still in force and that the speaker believes what she has said. Jokes and fictions might be seen as cases in which the maxim of truthfulness is overtly *suspended* (the speaker overtly *opts out* of it); the hearer is meant to notice that it is no longer operative, and is not expected to assume that the speaker believes what she has said. Metaphor, irony and other tropes represent a third category: they are *overt violations* (*floutings*) of the maxim of truthfulness, in which the hearer is meant to assume that the maxim of truthfulness is no longer operative, but that the supermaxim of Quality remains in force, so that some true proposition is still conveyed.

We will grant that a reasonable—if not optimal—treatment of lies, jokes and fictions might be developed along these lines. Tropes, and more generally loose uses of language, present a much more pressing challenge. After all, many, if not most, of our serious declarative utterances are not strictly and literally true, either because they are figurative, or simply because we express ourselves loosely.

2 The case of tropes

Lewis (1975; 1983) considers the case of tropes:

> *Objection*: Suppose the members of a population are untruthful in their language £ more often than not, not because they lie, but because they go in heavily for irony, metaphor, hyperbole, and such. It is hard to deny that the language £ is used by such a population.
> *Reply*: I claim that these people are truthful in their language £, though they are not *literally truthful* in £. To be literally truthful in £ is to be truthful in

> another language related to £, a language we can call literal-£. The relation between £ and literal-£ is as follows: a good way to describe £ is to start by specifying literal-£ and then to describe £ as obtained by certain systematic departures from literal-£. This two-stage specification of £ by way of literal-£ may turn out to be much simpler than any direct specification of £. (Lewis 1983: 183)

Lewis's reply rests on a widely shared view, which dates back to classical rhetoric. On this view,

(2) (a) Figurative and literal utterances differ not in the kind of meanings they have (thus, if literal meanings are truth-conditional, so are figurative meanings), but in the way these meanings are expressed and retrieved.
 (b) The meanings of figurative utterances are derived by systematic departures from their literal meanings.

For example, consider (3) and (4), where (3) is a metaphor and (4) is intended as hyperbole[1]:

(3) The leaves danced in the breeze.
(4) You're a genius!

Lewis might want to say that in literal-English, sentences (3) and (4) have just their literal meanings. In actual English, the language in which a convention of truthfulness and trust obtains among English speakers, (3) and (4) are ambiguous. They have their literal meanings plus other, figurative meanings: thus, (3) has the metaphorical meaning in (5), and (4) the hyperbolical meaning in (6):

(5) The leaves moved in the breeze as if they were dancing.
(6) You're very clever!

However, it is not as if any language £ (in the sense required by Lewis, where the sentences of £ can be assigned truth-conditional meanings) had ever actually been specified on the basis of a corresponding "literal-£". So what reason is there for

[1] In this paper we will focus on metaphor, hyperbole and a range of related tropes which we analyse as varieties of loose talk. For analyses of irony and understatement, see Sperber & Wilson (1981; 1986, chapter 4, sections 7, 9; 1990; 1998b); Wilson & Sperber (1992).

accepting something like (2a) and (2b)? How, and under what conditions, are figurative meanings derived from literal meanings? Lewis does not explain, and there are no generally accepted answers to these questions. We have argued (Sperber & Wilson 1986a,b; 1990; 1998a) that figurative interpretations are radically context-dependent, and that the context is not fixed independently of the utterance but constructed as an integral part of the comprehension process. If so, then the very idea of generating the sentences of a language £ on the basis of a corresponding "literal-£" is misguided.

Grice is often seen as having provided an explanation of how figurative interpretations are conveyed. Consider a situation where the speaker of (3) or (4) manifestly could not have intended to commit herself to the truth of the propositions literally expressed: it is common knowledge that she knows that leaves never dance, or that she does not regard the hearer as a genius. She is therefore overtly violating the maxim of truthfulness: in Grice's terms, she is flouting it. Flouting a maxim indicates a speaker's intention: the speaker intends the hearer to retrieve an implicature that brings the full interpretation of the utterance (i.e., what is said plus what is implicated) as close as possible to satisfying the Co-operative Principle and maxims. In the case of tropes, the required implicature is related to what is said in one of several possible ways, each characteristic of a different trope. With metaphor, the implicature is a simile based on what is said; with irony, it is the opposite of what is said; with hyperbole, it is a weaker proposition, and with understatement, a stronger one. Thus, Grice might analyse (3) as implicating (5) above, and (4) as implicating (6).

Note that this treatment of tropes does not differ radically from Lewis's, or from the classical rhetorical account. Grice's approach, like Lewis's, is based on assumption (2a) and, more importantly, on assumption (2b) (that the meanings of figurative utterances are derived by systematic departures from their literal meanings). The only difference is that Lewis sees these departures as systematic enough to be analysed in code-like terms: the figurative meaning of a sentence is a genuine linguistic meaning specified in the grammar of £ by a derivation that takes the literal meaning of the sentence as input. The sentences of £ (unlike those of literal-£) are systematically ambiguous between literal and figurative senses. For Grice, by contrast, sentences have only literal meanings. Figurative meanings are not sentence meanings but utterance meanings, derived in a conversational context. However, the derivations proposed in Grice's pragmatic approach to tropes are the same as those hinted at by Lewis in his linguistic approach, and neither differs seriously from the classical rhetorical account.

Grice's treatment of tropes leaves several questions unanswered, and we will argue that it is inconsistent with the rationale of his own enterprise. In particular, there is room for doubt about what he meant by the maxim of truthfulness, and the role it was intended to play in his framework. This doubt is created by two possible interpretations of his notion of SAYING. On the first interpretation, *saying* is merely expressing a proposition, without any necessary commitment to its truth. Understood in this way, the maxim of truthfulness means "Do not express propositions you believe to be false." The function of this maxim, and more generally of the Quality maxims, would be to account for the fact—to the extent that it is a fact—that a speaker actually commits herself to the truth of what she says. Tropes would then be explained by the claim that flouting the maxim triggers the recovery of an implicature in the standard Gricean way. However, there is a problem. In general, the recovery of implicatures is meant to restore the assumption that the maxims have been observed, or that their violation was justified in the circumstances (as when a speaker is justified by her ignorance in providing less information than required) (Grice 1989: 370). In the case of tropes, the maxim of truthfulness is irretrievably violated, and the implicature provides no circumstantial justification whatsoever.

On the second, and stronger, interpretation, *saying* is not merely expressing a proposition but asserting it, i.e. committing oneself to its truth. Understood in this way, the maxim of truthfulness means "Do not *assert* propositions you believe to be false." On this interpretation, saying already involves speaker commitment, and the function of the maxim of truthfulness, and more generally of the Quality maxims, would be to ensure that speakers do not make spurious commitments. This seems to fit with Grice's above remark that the function of the Quality maxim is to guarantee that contributions are genuine rather than spurious. However, understood in this way, it is hard to see why a maxim of truthfulness is needed at all. It seems to follow from the very notion of an assertion as a commitment to truth (perhaps together with a general injunction to fulfil your commitments) that your assertions should be truthful. In fact, the only pragmatic function of the maxim of truthfulness, on this interpretation, is to be violated in metaphor and irony, thus triggering the search for an implicature. Without it, Grice would have no account of figurative utterances at all.

Which notion of saying did Grice have in mind in formulating the maxim of truthfulness? There is evidence of some hesitation. On the one hand, he treats the tropes as "Examples in which the first maxim of Quality is flouted" (Grice 1989: 34). On the other, he comments that in irony the speaker "has said *or has made as if to say*" [our italics] something she does not believe, and that in metaphor what is communicated must be obviously related to what the speaker "*has made as if to say*" (ibid.: 34). If the speaker of metaphor or irony merely "makes as if to say" something, then the stronger notion of saying must be in force; on the other hand, if the speaker of a trope merely “makes as if” to say something, then surely the maxim of truthfulness is not violated. But if the maxim of truthfulness is not violated, how does Grice's analysis of metaphor and irony go through at all?

Elsewhere in his philosophy of language, where the notion of saying plays a central role, it was the stronger rather than the weaker notion that interested Grice. He says, for example (Grice 1989: 87): "I want to say that (1) 'U (utterer) said that *p*' entails (2) 'U did something x by which U meant that *p*'" (Grice 1989: 87). For Grice, what is *meant* is roughly co-extensive with what is *intentionally communicated*, i.e. with the information put forward as true. On this interpretation, saying involves speaker commitment: i.e. it means asserting. Among his commentators, Stephen Neale (1992) treats these broader considerations as decisive, and concludes that metaphor and irony are not cases of saying at all: "If U utters the sentence 'Bill is an honest man' ironically, on Grice's account U will *not* have said that Bill is an honest man: U will have made as if to say that Bill is an honest man." (Neale 1992, section 2).

How can we reconcile these two claims: that metaphor and irony are deliberate violations of the maxim of truthfulness, and hence must *say* something, and that in metaphor and irony the speaker merely *makes as if to say* something? A possible answer would be to distinguish two phases in the utterance interpretation process. In the first, the utterance of a declarative sentence would provide prima facie evidence for the assumption that an assertion is being made. In the second, this assumption would be evaluated and accepted or rejected. In the case of metaphor and irony, this second phase would involve an argument of the following sort: if it is common ground that the utterer U doesn't believe *p*, then U cannot assert *p*; it is common ground that U doesn't believe *p*; hence, U hasn't asserted *p*. In this way, we get a consistent interpretation of the notion of saying, and we can see why Grice hesitates between "saying" and "making as if to say."

However, if this interpretation is correct, then a trope involves no real violation of the maxim of truthfulness at any stage: since the speaker was not saying *p*, she was not saying what she believed to be false. A flouting, so understood, is a mere appearance of violation. So why should it be necessary to retrieve an implicature in order to preserve the assumption that the maxims have been respected? The Gricean way to go (although Grice himself did not take this route) would be to argue that it is not the maxim of truthfulness but some other maxim that is being violated. Quite plausibly, the maxim of Relation ("Be relevant") is being violated, for how can you be relevant when you speak and say nothing? Surely the first maxim of Quantity ("Make your contribution as informative as is required") is being violated, for if nothing is said, no information is provided. The implicature could thus be seen as a way of providing a full interpretation of the utterance in which these maxims are respected.

The problem with this analysis of tropes (and with the alternative analysis on which floutings of the maxim of truthfulness are *genuine* violations) is that it leads to an interpretation of figurative utterances that irretrievably violates the Manner maxims. In classical rhetoric, a metaphor such as (3) or a hyperbole such as (4) is merely an indirect and decorative way of communicating the propositions in (5) or (6). This ornamental value might be seen as explaining the use of tropes, insofar as classical rhetoricians were interested in explanation at all. Quite sensibly, Grice does not appeal to ornamental value. His supermaxim of Manner is not "Be fancy" but "Be perspicuous." He assumes, this time in keeping with classical rhetoric, that figurative meanings, like literal meanings, are fully propositional, and always paraphrasable by means of a literal utterance. Which raises the following question: isn't a direct and literal expression of what you mean *always* more perspicuous (and in particular less obscure and less ambiguous, cf. the first and second Manner maxims) than an indirect figurative expression? (Remember: you cannot appeal to the subtle extra effects of tropes, since they are not considered, let alone explained, within the Gricean framework).

3 The case of loose talk

Tropes are the most striking examples of serious utterances where the speaker is manifestly not telling the strict and literal truth. Even more common are loose uses or rough approximations, as in (7)-(10):

(7) The lecture starts at five o'clock.
(8) Holland is flat.
(9) *Sue:* I must run to the bank before it closes.
(10) *Jane:* I have a terrible cold. I need a Kleenex.

These utterances are not strictly speaking true: lectures rarely start at exactly the appointed time, Holland is not a plane surface, Sue must hurry to the bank but not necessarily run there, and other brands of disposable tissue would do just as well for Jane. Such rough approximations are very widely produced and understood. Some are tied to a particular situation, produced once and then forgotten. Others may be regular and frequent enough to give rise to an extra sense, which may stabilise in an individual or a population: lexical broadening (along with lexical narrowing and metaphorical transfer) has been seen as one of the three main factors in semantic change (Lyons 1977, 13.4, 14.5). What concerns us here is not so much the outcome of these historical macro-processes as the nature of the individual micro-processes that underlie them, and we will largely abstract away from the question of when a word such as "flat", or "run", or "Kleenex" may be said to have acquired an extra stable sense (see Sperber & Wilson 1998a for some discussion).

How should we analyse loose uses such as those in (7)-(10)? Are they lapses, the result of sloppy speech or thought, accepted by hearers whose expectations have been reduced to realistic levels by repeated encounters with normal human failings? Is it reasonable to assume that there really is a convention or maxim of truthfulness, although speakers quite commonly fall short of strictly obeying it? As hearers, would we always—and as speakers, should we always—prefer the strictly true statements in (11)-(14) to the loose talk of (7)-(10)?

(11) The lecture starts at or shortly after five.
(12) Holland has no mountain and very few hills.
(13) I must go to the bank as fast as if I were running.
(14) I need a Kleenex or other disposable tissue.

Clearly not. In most circumstances, the hearer would not be misled by strictly untrue approximations such as (7)-(10), and their strictly true counterparts in (11)-(14) would not provide him with any more valuable information. Indeed, since these strictly true counterparts are typically longer, the shorter approximations may be preferable.

Loose talk presents few problems for speakers and hearers, who are rarely even aware of its occurrence; but it does raise a serious issue for any philosophy of language based on a maxim or convention of truthfulness.

In Grice's framework, although approximations such as (7)-(10) apparently violate either the maxim of truthfulness or the second maxim of Quality ("Have adequate evidence for what you say"), they do not really fit into any of the categories of violation listed above. They are not covert violations, designed to deceive the hearer into believing the proposition strictly and literally expressed. They are not like jokes or fictions, which suspend the maxims entirely. A Gricean might try to analyse them as floutings: overt violations (real or apparent), designed to trigger the search for a related implicature (here a hedged version of what was literally said or quasi-said). The problem is that loose uses are not generally *perceived* as violating the Quality maxims at all. In classical rhetoric, they were not treated as tropes involving the substitution of a figurative for a literal meaning. They do not have the striking quality that Grice associated with floutings, which he saw as resulting in figurative or quasi-figurative interpretations. Loose talk involves no *overt* violation, real or apparent; or at least it does not involve the degree of overtness in real or apparent violation that might trigger the search for an implicature. While we are all capable of realising on reflection that utterances such as (7)-(10) are not strictly and literally true, these departures from truthfulness pass unattended and undetected in the normal flow of conversation. Grice's framework thus leaves them unexplained.

Perhaps we should reconsider the apparent platitudes we started with. Maybe we should have said that as speakers, we expect what we say to be accepted as *approximately* true, and as hearers, we expect what we are told to be *approximately* true. But this is far too vague. Approximations differ both in kind and in degree, and their acceptability varies with content and context. There is no single scale on which the degrees of approximation in disparate statements such as (7)-(10) can be usefully compared. The same statement can be an acceptable approximation in one situation and not in another. Thus, suppose the speaker of (7) expects the lecture to start at some time between five o'clock and 5.10: then (7) would be an acceptable approximation to a student who has just asked whether the lecture starts at five or six o'clock, but not to a radio technician preparing to broadcast the lecture live. Moreover, as we will argue below, there are cases in which the notion of "degrees of approximation" does not really apply.

A convention of truthfulness and trust in a language (if there were one) could play a crucial role in explaining how speakers and hearers achieve the co-ordination necessary for successful communication. If all that they are entitled to is vague expectations of approximate truth, it is hard to see how a robust enough convention of truthfulness could ever be established. Still, this is the direction that David Lewis proposes to explore. He writes:

> When is a sentence true enough? […] this itself is a vague matter. More important for our present purposes, it is something that depends on context. What is true enough on one occasion is not true enough on another. The standards of precision in force are different from one conversation to another, and may change in the course of a single conversation. Austin's "France is hexagonal" is a good example of a sentence that is true enough for many contexts but not true enough for many others. (1983: 244-245)

Lewis would agree with us that "hexagonal" and "flat" are absolute terms, and that their vagueness is pragmatic rather than semantic; however, his analysis of pragmatically vague terms such as "flat" is very similar to his analysis of semantically vague terms such as "cool".[2] For Lewis, a semantically vague term has a range of possible sharp delineations, marking different cut-off points between, say, "cool" and "warm". "This is cool" may be true at some but not all delineations, and depending on our purposes, we may be willing or unwilling to assert it: hence its vagueness (1983: 228-229). On Lewis's account, semantically absolute but pragmatically vague terms are handled on

[2] On the analysis of "flat", see Unger 1975, chap. 2; for discussion, see Gross 1998.

similar lines, except that semantic delineations are replaced by contextually-determined standards of precision (so if "flat" were semantically rather than pragmatically vague, the analysis would not be very different). On this approach, “Holland is flat” would be true according to some fairly low standard of precision, but false given higher standards.

Semantic vagueness clearly exists ("bluish" and “flattish” are good examples); its analysis raises problems of its own, about which we have nothing to say here (see Keefe & Smith 1996; Williamson 1994). What we do want to argue against is the idea that loose talk can be treated as a pragmatic analogue of semantic vagueness. As we will show, there are many varieties of loose talk, not all of which can be satisfactorily handled by appeal to contextually-determined standards of precision. For the cases that cannot be handled on Lewis’s lines, an alternative analysis must be found. We will propose such an analysis, and argue that it generalizes straightforwardly to all varieties of loose talk (and indeed to all utterances, literal, loose, or figurative), making the appeal to standards of precision as a component of conversational competence unnecessary.

In fact, there are problems even in some cases where the appeal to standards of precision looks initially plausible. Consider a situation where (7) ("The lecture starts at five o’clock") would be accepted as true enough if the lecture started somewhere between five o’clock and ten past five. On the lines proposed by Lewis, it might then be claimed that here the standard of precision allows for a give or take of, say, fifteen minutes around the stated time. It ought then to follow that a hearer in the same situation, with the same standard of precision in force, should be equally willing to accept (7) as true enough if the lecture started somewhere between 4.50 and five o’clock. But there is an obvious asymmetry between the two cases. Intuitively, the reason is clear enough: the audience won't mind or feel misled if they get to the lecture a few minutes early, but they will if they get there a few minutes late, so the approximation is acceptable only in one direction. In a different situation – when the speaker is talking about the end of the lecture rather than the beginning, for example – there may be an asymmetry in the other direction. Again, the reason is intuitively clear: the audience won't mind or feel misled if they can get away a bit earlier than expected, but they will if they have to stay longer. It is hard to explain these obvious intuitions using regular standards of precision. It would be possible, of course, to try building the asymmetries into the standards of precision themselves, but then two different standards would have to be invoked to explain how (15) is quite naturally understood to mean something like (16):

(15) The lecture starts at five o’clock and ends at seven o’clock.

(16) The lecture starts at five o’clock or shortly after and ends at seven o’clock or shortly before.

This is clearly ad hoc. It would be better to find an alternative account of these asymmetries; but then the appeal to standards of precision may become redundant.

A more serious problem for Lewis is that in some cases, the appeal to standards of precision does not seem to work at all. Lewis's account works best when there is a continuum (or ordered series) of cases between the strict truth and the broadest possible approximation. "Flat" is a good example, since departures from strict flatness may vary in degree. “Five o'clock” also works well in this respect, since departures from exactness may vary in degree. But with “run” in (9) ("I must run to the bank") and “Kleenex” in (10) ("I need a Kleenex"), no such continuum exists. There is a sharp discontinuity between running (where both feet leave the ground at each step) and walking (where there is always at least one foot on the ground). Typically (though not necessarily), running is faster than walking, so that "run" can be loosely used, as in (9), to indicate the activity of going on foot (whether walking or running) at a speed more typical of running. But walking at different speeds is not equivalent to running relative to different standards of precision. Similarly, Kleenex is a brand name: other brands of disposable tissue are not Kleenex. The word “Kleenex” can be loosely used, as in (10), to indicate a range of tissues similar to Kleenex. But there is no continuum on which being similar enough to Kleenex amounts to actually being Kleenex relative to stricter or looser standards of precision. "Run", "Kleenex" and very many other words have sharp conceptual boundaries and no ordered series of successively broader extensions that might be picked out by raising or lowering some standard of precision. Yet these terms are often loosely used. This shows that looseness is a broader notion than pragmatic vagueness.

Again, for someone with no particular theoretical axe to grind, it is easy enough to say intuitively what is going on. Suppose you have an afternoon lecture but don't know when it is due to start. Someone tells you, "The lecture starts at five o'clock." From the literal content of the utterance, together with other premises drawn from background knowledge, you can derive a number of conclusions that matter to you: that you will not be free to do other things between five and seven o'clock, that you should leave the library no later than 4.45, that it will be too late to go shopping after the lecture, and so on. To say that these conclusions matter to you is to say that you can use them to derive still further non-trivial contextual implications, of a practical or theoretical nature. These initial conclusions are the main branches of a derivational tree with many further branches and sub-branches. All these direct and indirect conclusions would also have been derivable from the strictly true utterance "The lecture starts at or shortly after 5.00," but at the extra cost required to process a longer sentence and a more complex meaning. There are other conclusions—false ones this time—that would have been derivable from the approximation "The lecture starts at five o'clock" but not from its strictly true counterpart: that the lecture will have begun by 5.01, for instance. But you are unlikely to derive them. They don't matter, because they are derivational bare branches which yield no further non-trivial implications.

Suppose Peter and Mary, who are both rather out of shape, are discussing where to go on their next cycling holiday. Mary suggests Holland, adding: "Holland is flat." From the strictly false proposition that Holland is flat—just as easily as from the strictly true hedged proposition that Holland is approximately flat—Peter can derive the true conclusion that cycling in Holland would involve no mountain roads and would not be too demanding. Unlike the true hedged proposition, the false approximation also has false implications (that there are no hills at all in Holland, for instance). But it is unlikely that Peter would even contemplate inferring any of these.

Suppose Sue, chatting with friends in the street, looks at her watch and says, "I must run to the bank before it closes." Her friends will take her to mean that she must break off their chat and hurry to the bank. For them, that much information is worth deriving. Whether she will actually get to the bank by running, walking fast or a mixture of both is of no interest to them, and they will simply not attend to this aspect of the literal meaning of her utterance.

Suppose Jane and Jack are in the cinema waiting for the film to start. By saying, "I have a terrible cold. I need a Kleenex," Jane provides a premise from which her friend can infer that she wants to borrow a tissue to use in dealing with her cold. Her utterance also provides a premise from which Jack could infer the possibly false conclusion that given a choice, Jane would prefer Kleenex to other brands of tissue; but he is unlikely to draw such a conclusion in this context, since it doesn't matter to him (and in a context where it would matter, the utterance would be inappropriate as a case of loose talk).

As these examples show, hearers do not object to strictly false approximations as long as the conclusions they bother to derive from them are true. In fact, for reasons of economy of effort, they might prefer the terser approximations to their longer-winded but strictly true counterparts.

Anticipating the arguments of the next section, let us say that an utterance is RELEVANT when the hearer, given his cognitive dispositions and the context, is likely to derive some genuine knowledge from it (we will shortly elaborate on this). Someone interested in defending a maxim or convention of truthfulness might then suggest that expectations of relevance do play a role in comprehension, but in a strictly limited way. One might claim, for example, that while utterances in general create expectations of truthfulness, approximations alone create expectations of relevance, which have a role to play in the case of loose talk, but only there. This account (apart from being unparsimonious) raises the following problem. As noted above, while we are all capable of realising on reflection that an utterance was an approximation rather than a strictly literal truth, in the normal flow of discourse, approximations are simply not recognised as such. But in that case, how could loose talk and literal talk be approached and processed with different expectations?

Here is the answer. It is not just approximations but all utterances—literal, loose or figurative—that are approached with expectations of relevance rather than truthfulness. Sometimes, the only way of satisfying these expectations is to understand the utterance as literally true. But just as an utterance can be understood as an approximation without being recognised and categorised as such, so it can be literally understood without being recognised and categorised as such. We will argue that the same is true of tropes. Literal, loose, and figurative interpretations are arrived at in the same way, by constructing an interpretation that satisfies the hearer's expectations of relevance.

No special machinery is needed to explain the interpretation of loose talk. In particular, contextually-determined standards of precision play no role in the interpretation process. They do not help with cases such as "run", or "Kleenex", which are neither semantically nor pragmatically vague; and to appeal to them in analysing cases such as "flat" or "at five o'clock", which might be seen as involving a pragmatic form of vagueness, would be superfluous at best.

4 Relevance: theory

Grice's maxim of truthfulness was part of what might be called an inferential model of human communication. This contrasts with a more classical code model, which treats utterances as signals encoding the messages that speakers intend to convey. On the classical view, comprehension is achieved by decoding signals to obtain the associated messages. On the inferential view, utterances are not signals but pieces of evidence about the speaker's meaning, and comprehension is achieved by inferring this meaning from the evidence provided. An utterance is, of course, a linguistically coded piece of evidence, so that the comprehension process will involve an element of decoding. But the linguistically-encoded sentence meaning need not be identical to the speaker's meaning—and we would argue that it never is—, since it is likely to be ambiguous and incomplete in ways the speaker's meaning is not. On this approach, the linguistic meaning recovered by decoding is just one of the inputs to an inferential process which yields an interpretation of the speaker's meaning.

Grice, Lewis and others who have contributed to the development of an inferential approach to communication have tended to minimise the gap between sentence meaning and speaker's meaning; they treat the inference from sentence meaning to speaker's meaning as merely a matter of assigning referents to referring expressions, and perhaps of calculating implicatures. While the slack between sentence meaning and speaker's explicit meaning cannot be entirely eliminated, a maxim or convention of truthfulness, if it is obeyed, has the effect of reducing it to a minimum. But why should this be a priori desirable? Comprehension is a complex cognitive process. From a cognitive point of view, how much of the work is done by inference and how much by decoding depends on how efficient the inferential processes are. We have argued (Wilson & Sperber 1981; Sperber & Wilson 1986; 1998) that relevance-oriented inferential processes are efficient enough to allow for a much greater slack between sentence meaning and speaker's meaning than is generally assumed. Here, we summarise the theory briefly for purposes of the present discussion.

We characterise RELEVANCE as a property of inputs to cognitive processes that makes them worth processing. ("Relevance" is used in a technical sense which is not meant to capture any of the ordinary senses of the word). These inputs may be external stimuli (e.g. a smell, the sound of an utterance), or internal representations which may undergo further processing (e.g. the recognition of a smell, a memory, the linguistic decoding of an utterance). At each point in our cognitive lives, we have many more potential inputs available than we can actually process; for example, we perceive many more distal stimuli than we can attend to, and have many more memories than we can reactivate at a single time. Efficiency in cognition is largely a matter of allocating our processing resources so as to maximise cognitive benefits. This involves processing inputs that offer the best expected cost/benefit ratio at the time.

Here we will consider only one type of cognitive benefit: improvements in knowledge. This is plausibly the most important type of cognitive benefit. There may be others—improvements in memory or imagination, for example; but it could be argued that these are benefits only because they contribute indirectly to improvements in knowledge: better memory and imagination lead to better non-demonstrative inference, and therefore to better knowledge. In any case, for our present purposes, there is another important reason for equating cognitive benefits with improvements in knowledge.

In a situation where it is clear to both participants that the hearer's goal in attending to the speaker's utterances is not the improvement of knowledge—say, he just wants to be amused—, there is no reason why the speaker should be expected to tell the truth. Thus, one way of challenging the maxim or convention of truthfulness would be to start by questioning whether humans are much interested in truth (see, for instance, Stich 1990). Here, we want to present a more pointed challenge to Grice's and Lewis's ideas, based on the nature of human communication rather than the goals of cognition. We will therefore grant (and not just for the sake of argument, but because we accept it as roughly correct) that one of the goals of most human communication (though certainly not the only one) is the transmission of genuine information and the improvement of the hearer's knowledge. We will consider only cases where hearers are interested in truth. Our claim is that even in these cases, hearers do not expect what is said to be strictly and literally true.

The processing of an input in the context of existing assumptions may improve the individual's knowledge not only by adding a new piece of information, but by revising his existing assumptions, or by yielding conclusions not derivable from the new piece of knowledge alone or from existing assumptions alone. We define an input as relevant when and only when it has such positive cognitive effects. Relevance is also a matter of degree, and we want to characterise it as not only a classificatory but also a comparative notion. There are potential inputs with some low degree of relevance all around us, but mere relevance is not enough. What makes an input worth attending to is not that it is relevant, but that it is more relevant than any alternative potential input to the same processing resources at that time. Although relevance cannot be measured in absolute terms, the relevance of various inputs may be compared.

For our purpose—which is to characterise a property crucial to cognitive economy—it must be possible to compare the relevance of inputs not only in terms of benefits (i.e. positive cognitive effects), but also in terms of costs (i.e. processing effort). We therefore propose the following comparative notion of relevance:

(17) *Relevance of an input to an individual at a time*
  (a) Everything else being equal, the greater the positive cognitive effects achieved in an individual by processing an input at a given time, the greater the relevance of the input to that individual at that time.
  (b) Everything else being equal, the smaller the processing effort expended by the individual to achieve those effects, the greater the relevance of the input to that individual at that time.

Here is a brief and artificial illustration. Peter wakes up feeling unwell and goes to the doctor. On the basis of her examination, she might make any of the following true statements:

(18) You are ill.
(19) You have flu.
(20) You have flu or 29 is the square root of 843.

The literal content of all three utterances would be relevant to Peter. However, (19) would be more relevant than (18), since it allows him to derive all the consequences derivable from (18) and more besides. This is an application of clause (a) of the characterisation of relevance in (17). (19) would also be more relevant than (20), because although exactly the same consequences are derivable from both, (20) requires more processing effort (in order to realise that the second disjunct is false and the first is therefore true). This is an application of clause (b) of the characterisation of relevance in (17).

Given this characterisation of relevance, it is, *ceteris paribus*, in the individual's interest to process the most relevant inputs available. We claim that this is what people tend to do (with many failures, of course). They tend to do it not because they realise that it is in their interest (and they certainly do not realise it in those terms), but because they are cognitively-endowed evolved organisms. In biological evolution, there has been constant pressure on the human cognitive system to organise itself so as to select inputs on the basis of their expected relevance. Hence:

(21) *The First, or Cognitive, Principle of Relevance*
The human cognitive system tends towards processing the most relevant inputs available.

The tendency described in the cognitive principle of relevance is strong enough, and manifest enough, to make our mental processes somewhat predictable to others. We are in general fairly good at predicting which (if any) of the external stimuli currently affecting some other individual's nervous system she is likely to be attending to, and which of the indefinitely many inferences that she might draw from it she will in fact draw. What we do, essentially, is assume that she will pay attention to the potentially most relevant stimulus, and process it so as to maximise its relevance: that is, in a context of easily accessible background assumptions, where the information it provides will carry relatively rich cognitive effects.

This mutual predictability is exploited in communication. As communicators, we provide stimuli which are likely to strike our intended audience as relevant enough to be worth processing, and to be interpreted in the intended way. A communicator produces a stimulus—say an utterance—that attracts her audience's attention, and she does so in an overtly intentional way. In other words, she makes it manifest that she wants her audience's attention. Since it is also manifest that the audience will tend to pay appropriate attention only to an utterance that seems relevant enough, it is manifest that the communicator wants her audience to assume that the utterance is indeed relevant enough. There is thus a minimal level of relevance that the audience is encouraged to expect: the utterance should be relevant enough to be worth the effort needed for comprehension.

Is the audience entitled to expect more relevance than this? In certain conditions, yes. The communicator wants to be understood. An utterance is most likely to be understood when it simplifies the hearer's task by demanding as little effort from him as possible, and encourages him to pay it due attention by offering him as much effect as possible. The smaller the effort, and the greater the effect, the greater the relevance. It is therefore manifestly in the communicator's interest for the hearer to presume that the utterance is not just relevant enough to be worth his attention, but more relevant than this. How much more? Here, the communicator is manifestly limited by her own abilities (to provide appropriate information, and to present it in the most efficient way). Nor can she be expected to go against her own preferences (against the goal she wants to achieve in communicating, for instance, or the rules of etiquette she wishes to follow). Still, it may be compatible with the communicator's abilities and preferences to go beyond the minimally necessary level of relevance. We define a notion of *optimal relevance* (of an utterance, to an audience) which takes these ideas into account, and propose a second principle of relevance based on it:

(22) *Optimal relevance of an utterance*
An utterance is optimally relevant to the hearer iff:
(a) It is relevant enough to be worth the hearer's processing effort;
(b) It is the most relevant one compatible with the speaker's abilities and preferences.

(23) *The Second, or Communicative, Principle of Relevance*
Every utterance conveys a presumption of its own optimal relevance.

In interpreting an utterance, the hearer invariably has to go beyond the linguistically-encoded sentence meaning. There will be ambiguities and referential indeterminacies to resolve, and other underdeterminate aspects of explicit content that we will look at shortly. There may be implicatures to identify. Achieving all this depends on the choice of an appropriate context. The communicative principle of relevance justifies the use of the following interpretation procedure: the hearer should consider interpretive hypotheses (disambiguations, reference assignments, implicatures, etc.) in order of accessibility—i.e. follow a path of least effort—and stop when he arrives at an interpretation that satisfies the expectations of relevance raised by the utterance itself.[3]

[3] A hearer's expectations of relevance may be more or less sophisticated. 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 appear relevant, what is expected is a speaker's meaning that it may have seemed to the speaker would

What makes it reasonable for the hearer to follow a path of least effort is that the speaker is expected (within the limits of her abilities and preferences) to formulate her utterance in such a way as to diminish the hearer's effort. Thus, the very fact that an interpretive hypothesis is easily accessible increases its initial plausibility (an epistemic advantage specific to communicated information).

What makes it reasonable for the hearer to stop at the first interpretation that satisfies his expectations of relevance is that there can never be more than one. A speaker who expressed herself in such a way that two or more interpretations yielded the expected level of cognitive effect would cause the hearer the gratuitous and unexpected extra effort of choosing among them, and the resulting interpretation (if any) would not satisfy clause (b) of the presumption of optimal relevance. Thus, when a hearer following the path of least effort finds an interpretation that is relevant as expected, this is the best possible interpretive hypothesis in the absence of contrary evidence. Since comprehension is a non-demonstrative inference process, this hypothesis may well be false. Typically, this happens when the speaker expresses herself in a way that is inconsistent with the expectations she herself has raised, so that the normal inferential routines of comprehension fail. Failures in communication are common enough: what is remarkable and calls for explanation is that communication works at all.

5 Relevance: illustration

An utterance has two immediate effects: it indicates that the speaker has something to communicate, and it determines an order of accessibility in which interpretive hypotheses will occur to the hearer. Here is an illustration.

Lisa drops by her neighbours, the Joneses, who have just sat down to supper:

(24) *Alan Jones*: Do you want to join us for supper?
Lisa: No, thanks. I've eaten.

A standard semantic analysis of the second part of Lisa's utterance would assign it the following truth condition:

(25) At some point in a time span whose endpoint is the time of utterance, Lisa has eaten something.

---

seem optimally relevant to the hearer. Adult communicators may nevertheless expect actual optimal relevance by default. Here we will ignore these complexities, but see Sperber (1994); Wilson (2000).

Clearly, though, Lisa means something more specific than this. She means that she has eaten that very evening, and not just anything, but a supper or something equivalent: a few peanuts wouldn't do.

Here is our explanation of how Alan understands Lisa's meaning. Her utterance activates in his mind, via automatic linguistic decoding, a conceptual structure that articulates in the grammatically specified way the concepts of Lisa, of eating, and of a time span whose endpoint is the time of utterance. He does not have to reason, because it is all routine, but he might reason along the following lines: she has caused me a certain amount of processing effort (the effort involved in attending to her utterance and decoding it). Given the communicative principle of relevance, this effort was presumably not caused in vain. So the conceptual structure activated by her utterance must be a good starting point for inferring her meaning, which should be relevant to me.

Lisa's utterance "I have eaten" immediately follows her refusal of Alan's invitation to supper. It would be relevant to Alan (or so she may have thought) to know the reasons for her refusal, which have implications for their relationship: Did she object to the offer? Would she accept it another time? It all depends on the reasons for her refusal. The use of the perfect "have eaten" indicates a time span with a definite endpoint—the time of utterance—and an indefinite starting point somewhere in the past. Alan narrows the time span by assuming that it started recently enough for the information that Lisa has eaten during that period to yield adequate consequences: here, the relevant time span is that very evening (Wilson & Sperber 1998). He does the same in deciding what she ate. In the circumstances, the idea of eating is most easily fleshed out as eating supper, and this, together with the narrowing of the time span, yields the expected level of cognitive effect. Alan then assumes that Lisa intended to express the proposition that she has eaten supper that evening, and to present this as her reason for refusing his invitation. This attribution of meaning is typically a conscious event; but Alan is never aware of the process by which he arrived at it, or of a literal meaning equivalent to (25).

This comprehension process can be represented in the form of a table, as in (26), with Alan's interpretive hypotheses on the left, and his basis for arriving at them on the right. We have presented the hypotheses in English, but for Alan they would be in whatever is the medium of conceptual thought, and they need not correspond very closely to our paraphrases:

(26)

(a) Lisa has said to Alan "I have eaten"	*Decoding of Lisa's utterance*
(b) Lisa's utterance is optimally relevant to Alan	*Expectation raised by the recognition of Lisa's utterance as a communicative act, and the acceptance of the presumption of relevance it automatically conveys*
(c) Lisa's utterance will achieve relevance by explaining her immediately preceding refusal of Alan's invitation to supper	*Expectation raised by (b), together with the fact that such an explanation would be most relevant to Alan at this point*
(d) The fact that one has already eaten supper on a given evening is a good reason for refusing an invitation to supper that evening	*First assumption to occur to Alan which, together with other appropriate premises, might satisfy expectation (c). Accepted as an implicit premise of Lisa's utterance*
(e) Lisa has eaten supper this evening	*First enriched interpretation of Lisa's utterance as decoded in (a) to occur to Alan which might combine with (d) to lead to the satisfaction of (c). Accepted as Lisa's explicit meaning*
(f) Lisa is refusing supper with us because she has already had supper this evening	*Inferred from (d) and (e), satisfying (c) and accepted as an implicit conclusion of Lisa's utterance*
(g) Lisa might accept an invitation to supper another time	*From (f) plus background knowledge. One of several possible weak implicatures of Lisa's utterance which, together with (f), satisfy expectation (b)*

We do not see this comprehension process as sequential, starting with (a) and ending with (g). For one thing, interpretation is carried out "on line," and starts while the utterance is still in progress. Some tentative or incomplete interpretive hypotheses can be made and later revised or completed in the light of their apparent consequences for the overall interpretation. We assume, then, that interpretive hypotheses about explicit content and implicatures are developed in parallel, and stabilise when they are mutually adjusted, and jointly adjusted with expectations of relevance.

In the present case, Alan assumes in (b) that Lisa's utterance, decoded as in (a), is optimally relevant to him. Since what he wants to know at this point is why she refused his invitation, he assumes in (c) that her utterance will achieve relevance by answering this question. In this context, Lisa's utterance “I have eaten” provides easy access to the piece of common background knowledge in (d)—that people don't normally want to eat supper twice in one evening. This could be used as an implicit premise in deriving the expected explanation of Lisa's refusal, provided that her utterance is interpreted on the explicit side as conveying the information in (e): that she has eaten supper that evening. By combining the implicit premise (d) and the explicit premise (e), Alan arrives at the implicit conclusion (f), from which further weaker implicatures, including (g) and others, can be derived (on the notion of a weak implicature, see below). This overall interpretation satisfies Alan's expectations of relevance. On this account, explicit content and implicatures (implicit premises and conclusions) are arrived at by a process of mutual adjustment, with hypotheses about both being considered in order of accessibility.

There is a certain arbitrariness in the way we have presented Alan's interpretive hypotheses. This is partly because, as noted above, we had to put into English thoughts that may not have been articulated in English. Another reason is that Lisa's utterance licenses not a single interpretation but any one of a range of interpretations with very similar import. By constructing any particular interpretation from this range, the hearer achieves comprehension enough. Thus, Alan might understand Lisa as meaning either that she has had supper that evening or, more cautiously, that, whether or not what she has eaten can properly be described as "supper", she has eaten enough not to want supper now. He may take her to be implicating (g), or some conclusion similar to (g), or nothing of the sort. In each case his interpretation is reasonable, in the sense that Lisa's utterance has encouraged him to construct it.

If Alan understands Lisa as meaning that she has had supper, or as implicating something like (g), he is taking a relatively greater share of the responsibility for his interpretation of her utterance. But this is something that hearers often do, and that speakers intend, or at least encourage, them to do. Often, the hearer will be unable to arrive at an interpretation that is relevant in the expected way without taking some of the responsibility for it: that is, without going beyond what the speaker commits herself to acknowledging as exactly what she meant. This is typical in loose talk and creative metaphor, both of which involve the communication of weak implicatures: implicatures that the hearer is given some encouragement but no clear mandate to construct. Nor is this sharing of responsibility a sign of poor communication: it may be just the degree of communication that suits both speaker and hearer.

Lisa's explicit meaning, as understood by Alan, logically implies the literal, unenriched meaning of her utterance: if she has eaten supper that evening, she has eaten *tout court*. From an analytical point of view, her utterance might therefore be classified as literal, for whatever good it might do. However, the literal meaning is not attended to at any stage, and the fact that the utterance is literal plays no role in the communication process. This is even more obvious in the following alternative version of the dialogue:

(27) *Alan*: Do you want to join us for supper?
Lisa: I'd love to. I haven't eaten.

Here, if the literal meaning of Lisa's utterance "I haven't eaten" is the negation of (25), i.e. the proposition that she has never eaten anything, then it is a blatant falsehood. However, this absurd interpretation never crosses the communicators' minds.

One way of avoiding such counterintuitive assignments of literal meaning would be to assume that the perfect 'has eaten' contains a hidden linguistic constituent which denotes a contextually determinate time span. In (27), Lisa could then be seen as referring, via this hidden constituent, to the evening of utterance, and the fact that she has eaten plenty in her lifetime would not falsify her statement, even literally understood. We will argue below that this move is ad hoc and unnecessary, but let us accept it here for the sake of argument.

Assume, then, that the literal meaning of Lisa's utterance in (27) is that she has not eaten anything that evening. Now suppose that she has in fact eaten a couple of peanuts, so that her utterance is strictly speaking false. Though it may be false, it is not misleading. Rather, it is a case of loose talk. Alan understands Lisa as saying that she has not eaten supper that evening. He arrives at this interpretation by taking the concept of eating, activated in his mind by automatic linguistic decoding, and narrowing it down to the concept of eating supper, which yields an interpretation that is relevant in the expected way. The procedure is the same as for dialogue (24), but since the narrowed concept falls within the scope of a negation, the resulting overall interpretation involves a loosening rather than a narrowing of the literal meaning.

It could be argued, of course, that Lisa's utterance contains a second hidden linguistic constituent, which denotes the food she has eaten. On this interpretation, the linguistically-determined truth-conditional meaning of "I have eaten" is equivalent not to "I have eaten something", but to "I have eaten x", where the value of "x" (like the referent of the pronoun "I" and the time of utterance) must be specified before the sentence token can be said to express a proposition.

In other situations, what the speaker means by saying that she has or hasn't eaten might also involve a specification of the place of eating, some manner of eating, etc:

(28) "I've often been to their parties, but I've never eaten anything" [*there*]
(29) "I must wash my hands: I've eaten" [*using my hands (rather than, say, being spoon-fed)*]

To deal with all such cases, more and more hidden constituents could be postulated, so that every sentence would come with a host of hidden constituents, ready for all kinds of ordinary or extraordinary pragmatic circumstances. In this way, the very idea of loose talk could be altogether avoided. We see this as a *reductio* argument that goes all the way to challenging what we accepted earlier for the sake of argument: that the use of the perfect carries with it a hidden constituent denoting a given time span. There is no need to postulate such a hidden constituent: the same process that explains how "eating" is narrowed down to "eating supper" also explains how the time span indicated by the perfect is narrowed down to the evening of utterance. Moreover, the postulation of such hidden constituents is ad hoc: its role is to reduce to a minimum the slack between sentence meaning and speaker's meaning, a slack that is uncomfortable from certain theoretical viewpoints. However, we read the evidence as showing that the slack actually *is* considerable, and we adopt a theoretical viewpoint that might help us describe and understand this fact. (For further discussion, see Carston 2000.)

6 The explicit communication of unencoded meanings

We are exploring the idea that the linguistically-encoded sentence meaning gives no more than a schematic indication of the speaker's meaning. The hearer's task is to use this indication, together with background knowledge, to construct an interpretation of the speaker's meaning, guided by expectations of relevance raised by the utterance itself. The conceptual resources brought to this task include all the concepts encoded in the hearer's language, but they go well beyond this (Sperber &Wilson 1998a). In particular, a concept may be recognised in context as a constituent of the speaker's explicit[4]

[4] This obviously involves rethinking the notion of explicitness itself. We do this in the final section.

meaning even though there is no expression in the sentence uttered, or indeed the language, which has this concept as its linguistically-encoded meaning. This happens regularly in loose talk.

Consider Sue chatting to her friend Jim in the street. Suddenly, she looks at her watch and says:

(30) I can't stay. I must run to the bank.

The process by which Jim interprets Sue's utterance "I must run to the bank" can be represented, as before, in table form:

(31)

(a) Sue has said to Jim, "I must run to the bank."	*Decoding of Sue's utterance*
(b) Sue's utterance is optimally relevant to Jim	*Expectation raised by the recognition of Sue's utterance as a communicative act, and the acceptance of the presumption of relevance it automatically conveys*
(c) Sue's utterance will achieve relevance by explaining why she must break off their chat	*Expectation raised by (b), together with the fact that such an explanation would be most relevant to Jim at this point*
(d) Having to hurry to the bank on urgent business is a good reason for breaking off a chat	*First assumption to occur to Jim which, together with other appropriate premises, might satisfy expectation (c). Accepted as an implicit premise of Sue's utterance*
(e) Sue must RUN* to the bank (where RUN* is the meaning indicated by "run", and is such that Sue's having to RUN* to the bank is relevant-as-expected in the context)	*(Description of) the first enriched interpretation of Sue's utterance as decoded in (a) to occur to Jim which might combine with (d) to lead to the satisfaction of (c). Interpretation accepted as Sue's explicit meaning*
(f) Sue must break off their chat because she must hurry to the bank on urgent business	*Inferred from (d) and (e), satisfying (c) and accepted as an implicit conclusion of Sue's utterance*
(g) Sue is afraid that if she stays chatting any longer, the bank may close before she gets there	*From (f) plus background knowledge. One of several possible weak implicatures of Sue's utterance which, together with (f), satisfy expectation (b)*

What Jim takes to be Sue's explicit meaning is describable as (31e):

(31e) Sue must RUN* to the bank (where RUN* is the meaning indicated by "run", and is such that Sue's having to RUN* to the bank is relevant-as-expected in the context).

This is not, of course, a proper paraphrase—let alone a proper analysis—of Sue's meaning (as understood by Jim). The notions of *a meaning indicated by a word* and of *relevance-as-expected in a context* are not constituents of Sue's meaning, and Jim does not have to deploy them in order to understand Sue. As it stands, (31e) is not an interpretation but merely a description of Sue's meaning. It attributes to Sue's utterance the property of indicating rather than encoding her meaning, and to Sue's meaning the property of justifying the derivation of enough cognitive effects to make her utterance worth Jim's processing effort. However, it goes without saying that if Jim is successful at all in understanding Sue's utterance, the result of his interpretation process will be not a description but an interpretation of Sue's meaning: that is, a mental representation which, if not identical to Sue's meaning, is similar enough in content to count as comprehension. In particular, Jim's interpretation must contain an unglossed version of the concept RUN*, which on our account was not encoded, but merely indicated, by her use of the word "run".

Let us assume that a satisfactory account can be provided of the nature of these concepts and of how hearers may grasp them (we will shortly return to this). Then an analysis on the lines of (31) shows how a term like "run", or "Kleenex"—which is neither semantically nor pragmatically vague, and which cannot be dealt with by appeal to contextually-determined standards of precision along Lewisian lines—can be loosely used and understood. As we will show, the analysis is straightforwardly generalisable to the full range of cases, including "flat" and "five o'clock", making the appeal to contextually-determined standards of precision unnecessary.

Consider Peter and Mary discussing their next cycling trip. Peter has just said that he feels out of shape. Mary says:

(32) We could go to Holland. Holland is flat.

The process by which Peter interprets Mary's utterance "Holland is flat" can be schematically represented, as before, in table form:

(33)

(a) Mary has said to Peter, "Holland is flat"	*Decoding of Mary's utterance*
(b) Mary's utterance is optimally relevant to Peter	*Expectation raised by the recognition of Mary's utterance as a communicative act, and the acceptance of the presumption of relevance it automatically conveys*
(c) Mary's utterance will achieve relevance by giving reasons for her proposal to go cycling in Holland, which take account of Peter's immediately preceding complaint that he feels out of shape	*Expectation raised by (b) together with the fact that such reasons would be most relevant to Peter at this point*
(d) Cycling on relatively flatter terrain which involves little or no climbing is less strenuous and would be agreeable in the circumstances	*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*
(e) Holland is FLAT* (where FLAT* is the meaning indicated by "flat", and is such that Holland's being FLAT* is relevant-as-expected in the context)	*(Description of) the first enriched interpretation of Mary's utterance as decoded in (a) to occur to Peter which might combine with (d) to lead to the satisfaction of (c). Interpretation accepted as Mary's explicit meaning*
(f) Cycling in Holland would involve little or no climbing	*Inferred from (d) and (e). Accepted as an implicit conclusion of Mary's utterance*
(g) Cycling in Holland would be less strenuous and would be agreeable in the circumstances	*Inferred from (d) and (f), satisfying (b) and (c) and accepted as an implicit conclusion of Mary's utterance*

What Peter takes to be Mary's explicit meaning is describable as (33e):

(33e) Holland is FLAT* (where FLAT* is the meaning indicated by "flat", and is such that Holland's being FLAT* is relevant-as-expected in the context).

As noted above, this is not an interpretation but merely a description of Mary's meaning. It attributes to Mary's utterance the property of indicating rather than encoding her meaning, and to Mary's meaning the property of warranting the derivation of enough cognitive effects to make her utterance worth processing for Peter. However, the outcome of Peter's comprehension process must be an interpretation rather than a description of Mary's meaning. In particular, it must contain an unglossed version of the concept FLAT*, which on our account was not encoded, but merely indicated, by her use of the word "flat".

First, what might this concept FLAT* be? It is not too difficult to give a rough answer. As Mary means it, a terrain is FLAT* if it can be travelled with little or no climbing. Being FLAT* is quite compatible with small-scale unevenness, and indeed with being not plane but convex because of the curvature of the Earth. (In fact, for a large country, being FLAT* is incompatible with being flat: if Holland were flat, travelling from the centre to the borders would involve going upwards, i.e. further away from the centre of the Earth.) The concept FLAT* indicated by Mary's utterance is even more specific. It has to do with cycling when not in great shape, which determines what will count as cases of climbing. On another occasion, when travelling by car and hoping to see mountain scenery, Mary might describe the south of England, say, as "flat"; however, what she would then mean is not FLAT* but some other concept which would be appropriately indicated, in this different context, by her use of the word "flat".

Next, how does Peter grasp the concept FLAT* indicated by Mary's utterance? We claim that, in appropriate circumstances, the relevance-theoretic comprehension procedure automatically guides the hearer to an acceptably close version of the concept conveyed. As noted above, the hearer's expectations of relevance warrant the assumption that the speaker's explicit meaning will contextually imply a range of specific consequences (made easily accessible, though not yet implied, by the linguistically-encoded sentence meaning). Having identified these consequences, he may then, by a process of backwards inference, enrich his interpretation of the speaker's explicit meaning to a point where it does carry these implications.

The claim that Holland is FLAT* carries a range of implications which Mary expects to fulfil Peter's expectations of relevance. The concept FLAT* is individuated—though not, of course, defined—by the fact that, in the situation described, it is the first concept to occur to Peter that determines these implications. If Mary is right about the implications that Peter will actually draw from her utterance, he should arrive by a process of spontaneous backwards inference at an appropriate understanding of her explicit meaning, and in particular of the concept FLAT*.

The implications that Mary expects Peter to derive need not be individually represented and jointly listed in her mind. In normal circumstances, they would not be. She might merely expect him to derive some implications which provide reasons for going cycling in Holland, and are similar in tenor to those she herself has in mind (again without necessarily having a distinct awareness of each and every one of them). To the extent that her expectations about the implications Peter will derive are indeterminate, the same will go for the concept she intends him to arrive at by backwards inference from these implications. Note that a difference in implications need not lead to a difference in concepts: from a somewhat different set of implications than the one envisaged by Mary, Peter may actually arrive at the same concept FLAT* that she had in mind. Suppose, though, that Peter constructs some concept FLAT**, which differs slightly from FLAT* but has roughly the same import in the situation. This would not be a case of imperfect communication or insufficient understanding. As noted above, it is quite normal for communicators to aim at such a relatively loose fit between speaker's meaning and hearer's interpretation.

We have described Mary's remark that Holland is flat as a case of loose talk. We could also have described it as a case of hyperbole, i.e. of a trope. After all, if it were taken literally, it would be a gross exaggeration. Nothing of substance hinges on whether Mary's utterance is categorised in one way or the other. The very same process of interpretation gives rise to literal, loose, hyperbolical or metaphorical interpretations, and there is a continuum of cases which cross-cut these categories.

Consider again the case of Peter and Mary discussing their next cycling trip. Peter has just said that he feels out of shape. In this version, Mary says:

(34) We could go to Holland. Holland is a picnic.

This is a clearly metaphorical use of "picnic". As before, Peter's interpretation of Mary's utterance can be represented in table form:

(35)

(a) Mary has said to Peter, "Holland is a picnic"	*Decoding of Mary's utterance*
(b) Mary's utterance is optimally relevant to Peter	*Expectation raised by the recognition of Mary's utterance as a communicative act, and the acceptance of the presumption of relevance it automatically conveys*
(c) Mary's utterance will achieve relevance by giving reasons for her proposal to go cycling in Holland, which take into account Peter's immediately preceding complaint that he feels out of shape	*Expectation raised by (b), together with the fact that such reasons would be most relevant to Peter at this point*
(d) Going on a picnic takes little effort	*First assumptions to occur to Peter which, together with other appropriate premises, might satisfy expectation (c). Accepted as implicit premises of Mary's utterance*
(e) Going on a picnic is a pleasant and relaxed affair	
(f) Holland is a PICNIC* (where PICNIC* is the meaning indicated by "picnic", and is such that Holland's being a PICNIC* is relevant-as-expected in the context)	*(Description of) the first enriched interpretation of Mary's utterance as decoded in (a) to occur to Peter which, together with (d) and (e), might lead to the satisfaction of (c). Interpretation accepted as Mary's explicit meaning*
(g) Going to Holland would involve little effort	*Inferred from (d) and (f), contributing to the satisfaction of (b) and (c), and accepted as an implicit conclusion of Mary's utterance*
(h) Going to Holland would be a pleasant and relaxed affair	*Inferred from (e) and (f), contributing to the satisfaction of (b) and (c), and accepted as an implicit conclusion of Mary's utterance*

Mary uses the word "picnic" to indicate the concept PICNIC*, which is part of what she wants to convey. Peter reconstructs this concept by treating the word "picnic", and its associated mental encyclopaedic entry, as a source of potential implicit premises such as (35d) and (35e). From these implicit premises and a still incomplete interpretation of Mary's explicit meaning, he tentatively derives implicit conclusions (35g) and (35h), which make the utterance relevant-as-expected in the situation. He then arrives by backwards inference at the full interpretation of the explicit content (35f), and its constituent concept PICNIC*. There is an unavoidable arbitrariness about the implicit premises and conclusions we have listed in table (35). The more metaphorical the interpretation, the greater the responsibility of the hearer in constructing implicatures (implicit premises and implicit conclusions), and the weaker most of these implicatures will be. In the case of more poetic metaphors, there is typically a wide range of potential implicatures, and the hearer or reader is encouraged to be creative in exploring this range (a fact well recognised in literary theory since the Romantics). Lack of an exact match between the implicatures envisaged by the speaker and those constructed by the hearer does not imply any failure of communication. The kind of communication aimed at in speaking metaphorically allows for, and indeed encourages, some freedom of interpretation.

The concepts FLAT* and PICNIC* referred to above are neither encoded nor encodable in English as spoken by Mary and Peter at the time of their exchange. No single word or phrase of English has these concepts as one of its linguistically-encoded senses. However, once Mary and Peter have successfully communicated such a concept, they may be able to co-ordinate more or less tacitly and adopt a new word or phrase to encode it (or add to the ambiguity of an existing word, for instance by giving the word "flat" another stable sense, namely FLAT*).

Different degrees of difficulty are involved in entertaining a linguistically unencodable concept such as FLAT* or PICNIC*, communicating such a concept, and lexicalising it. For any individual, entertaining a currently unencodable concept (i.e. a concept unencodable with the resources of the language at the time) is quite an easy and ordinary cognitive practice. We engage in such a practice every time we discriminate and think about a property for which there is no word or phrase in our public language. We may well do this several times a day. Communicating a content which has such an unencodable concept as a constituent is a matter of co-ordinating the cognitive activity of two people at a given point in time, causing them to pay attention to the same property or object; this is more difficult than doing it singly, but is still quite a common occurrence. Stabilising a word in the public language to express such a concept involves co-ordinating cognitive dispositions in a community over time; this is much more difficult and does not normally occur more than, say, a few times a year in a homogeneous speech community (see Sperber & Wilson 1998a).

7 Rethinking "explicit," "literal," and "what is said"

If the above analysis is correct, the notions *explicit*, *literal* and *what is said*, which Grice and Lewis saw as relatively unproblematic, will have to be rethought. For Grice, what is communicated by an utterance has two components: WHAT IS SAID and (optionally) WHAT IS IMPLICATED. He coined the terms "implicate" and "implicature" to refer to what is implicitly communicated, but rather than use the symmetrical "explicate" and "explicature", or just talk of "what is explicit," he chose to contrast what is implicated with the ordinary-language notion *what is said*. This terminological choice reflected both a presupposition and a goal. The presupposition was that *what is said* is an intuitively clear, common-sense notion. The goal was to argue against the view of meaning that ordinary-language philosophers were defending at the time. To achieve this goal, Grice wanted to show that *what is said* is best described in terms of a parsimonious semantics, while much of the complexity and subtlety of utterance interpretation is best explained in terms of implicatures (see Carston 1998, forthcoming for detailed discussion). We share Grice's goal—to relieve the semantics of natural language of whatever can be best explained at the pragmatic level—; but we take a substantially different view of how this pragmatic explanation should go.

Our account of linguistic communication assigns a theoretical status to the notions of EXPLICATURE and IMPLICATURE (roughly, the explicit and implicit contents of utterances). By contrast, it gives no theoretical status to the notions of *literal meaning*, or *what is said*. In fact, we coin the term "explicature", on the model of Grice's "implicature", because we doubt that there is any reliable common-sense notion of *what is said*. The explicature[5] of an utterance is partly decoded and partly inferred, while implicatures are wholly inferred. Inferring the explicature is a matter of disambiguating, enriching, and fine-tuning the semantic schema obtained by linguistic decoding. Inferring the implicatures is a matter of identifying implicit premises and conclusions which yield an interpretation that is relevant in the expected way. As we have shown, explicatures and implicatures are typically inferred in parallel, via mutual adjustment of interpretive hypotheses guided by considerations of relevance.

[5] We are talking here only of what we call first-level explicatures. We also claim that there are higher-level explicatures which normally do not contribute to the truth conditions of the utterance. For discussion, see Wilson & Sperber 1993.

We have already argued that implicatures can be stronger or weaker. The same is true of explicatures. Identifying the explicature of an utterance involves a certain amount of inference. The inference is non-demonstrative, and draws on background knowledge, so the hearer must always take some degree of responsibility for how it comes out. How much responsibility he has to take, and hence how indeterminate the explicature is, varies from utterance to utterance. Explicatures can be weaker or stronger, depending on the degree of indeterminacy introduced by the inferential aspect of comprehension. To illustrate, consider again dialogue (20), and three new versions of Lisa's answer in (36a-c):

(20) (a) *Alan Jones*: Do you want to join us for supper?
(b) *Lisa:* No, thanks. I've eaten.
(36) (a) *Lisa*: No, thanks. I've already eaten supper.
(b) *Lisa*: No, thanks. I've already eaten tonight.
(c) *Lisa*: No, thanks. I've already eaten supper tonight.

In understanding the explicit content of all four answers, a certain amount of inference, and hence a certain degree of indeterminacy, is involved. It might be thought that the only inferences involved in (36c) are automatic—simply a matter of fixing the referents of "I" and "tonight"—, but this would be a mistake. When Lisa describes what she has eaten as "supper," she may be speaking loosely. She may have eaten a sandwich, and be unwilling to have supper for that reason. So Alan might reasonably take "supper" to mean SUPPER*: that is, let us say, enough food to take the place of supper. If he understands Lisa to mean not SUPPER* but SUPPER (i.e. a regular evening meal), this is no less inferential. Whichever of the two interpretations is the first to come to mind will yield an overall interpretation that is relevant-as-expected, and will therefore be accepted.

Note that the encoded meaning SUPPER need not be the first to occur to Alan. Suppose he knows that Lisa generally has a salad or a sandwich instead of supper: then by saying that she has eaten "supper", she may make SUPPER* more immediately accessible than SUPPER. Generally speaking, encoded lexical senses need not, in the situation of utterance, be the most accessible ones. So when an encoded lexical sense is in fact chosen, the process is the same as when a word is understood to contribute a non-encoded sense. In each case, the first sense to be accessed and found to contribute to a relevant-as-expected interpretation is taken to be the intended one.

All four answers (20b) and (36a-c) communicate not just the same overall content, but also the same explicature and implicatures. If this is not immediately obvious, there is a standard test for deciding whether some part of what is communicated is explicitly or implicitly conveyed. This involves checking whether it falls within the scope of logical operators such as negation (see e.g. Wilson & Sperber 1998). Implicatures don't fall within the scope of negation and other operators, while explicatures do. Thus, consider the hypothesis that the explicature of (20b) is just the truism that Lisa has eaten at some point before the time of utterance, and that she is merely *implicating* that she has eaten that evening. This hypothesis does not survive the standard test. If Lisa had said "I haven't eaten" (as in dialogue (23)), she would clearly not have been saying that she has never eaten in her life, but merely denying that she has eaten "supper" (interpreted as SUPPER or SUPPER*) that very evening. So in saying "I've eaten," Lisa is explicitly communicating that she has eaten "supper" that very evening.

Though the same explicature—that Lisa has eaten SUPPER (or SUPPER*) that evening—is arrived at inferentially in all four cases, there is a clear sense in which it is weaker in (20b) than in (36a) or (36b), and stronger in (36c). The strength of an explicature varies with the degree of indeterminacy resulting from the inferential aspect of the recovery of explicatures. In particular, *ceteris paribus*, the more distant from any of its lexicalised senses is the concept conveyed by use of a word in a given utterance, the weaker the explicature will be. With metaphors, explicatures are at their weakest.

When the explicature of an utterance is quite strong, and in particular when the words are used to convey (one of) their encoded senses, what the theory describes as the explicature of the utterance corresponds closely to what might be common-sensically described as its "explicit content," or "what is said," or the "literal meaning" of the utterance. Whether the explicature is strong or weak, the notion of explicature applies unproblematically. However, the same is not true with the common-sense notions of *literal meaning* and *what is said*.

The notion of literal meaning, which plays such a central role in most theories of language use, is unclear in many respects. If literal meaning is defined as a linguistically-encoded sense of a sentence, then the literal meaning of a sentence never coincides with the speaker's explicit meaning in uttering this sentence (except in the case of genuine "eternal sentences," if any exist and are ever used). A speaker's meaning is typically propositional, and at the very least, reference has to be fixed in order to get from a sentence meaning to a proposition. The literal meaning of an utterance might then be defined as the proposition determined by combining linguistic sense with reference. When the speaker's meaning coincides with this proposition, we do indeed have a prototypical case of literalness. Imagine an anthropologist confessing:

(37) I have eaten human flesh.

In most situations, (37) would be relevant enough if its literal meaning were taken to be its explicature, without any narrowing of the time-span or the manner in which the eating of human flesh is understood to have taken place. This, then, is a prototypical case where literal meaning (understood as sense plus reference) and speaker's explicit meaning, or explicature, coincide. However, such cases are the exception rather than the rule.

First, there are cases where the explicature could not simply correspond to linguistic sense plus reference, because this is not enough to yield a proposition. Consider (38):

(38) His car is too big

Even by fixing the referents of "his" and the present tense, and combining the result with the linguistic sense, no propositional meaning is obtained.[6] "His car" could be the car he owns, the car he is renting, the car he is thinking about, and so on, and deciding which it is meant to be is not a matter of fixing reference, or disambiguating, but of enriching the linguistically-encoded meaning. Similarly, "too big" is indeterminate unless some contextual criterion is supplied for deciding what counts as big enough in this case. It might then be claimed that the literal meaning of an utterance is determined by a combination of literal sense, plus reference, plus obligatory enrichment (sometimes known as the "minimal proposition" expressed by an utterance; cf Bach 1994a,b; Carston 1988, 1998, forthcoming; Recanati 1989; Travis 1985). Suppose (38) is enriched to mean (39):

(39) The car Bob is planning to steal is too big to be hidden in the lorry.

Is this the literal meaning of (38) on that occasion, or is there some other, simpler literal meaning? If so, what is it? In such cases, intuitions about literalness become quite hazy.

Even leaving aside the problem raised by obligatory enrichment (and other related problems discussed in Searle 1979, ch. 5), and considering only sentences for which the combination of sense and reference yields a full proposition, the fact is that in most cases, the explicature of an utterance goes beyond this putative literal meaning. Identification of explicatures may involve enrichment of linguistic meaning, loosening of linguistic meaning, or a combination of both. Such cases are sometimes handled by

---

[6] Whatever the propositional meaning of a literal utterance of (38), it entails the existentially quantified proposition: *There is a relationship between the referent of "his" and a unique car, and there is a criterion of size, such that this car is too big by this criterion*. However, this proposition is never the utterance meaning of (38), and it would be highly counter-intuitive to claim that it is its literal meaning.

distinguishing *literal meaning* from *literal use*, and arguing that as long as the explicature departs from the literal meaning only by being richer, or more specific, then the utterance is literally used (see, for instance Katz 1990: 144-146). So when Lisa says "I have eaten" in response to Alan's invitation to supper, the literal meaning—the proposition that Lisa has eaten at some point in a time-span before the time of utterance—is determined in the regular way by a combination of sense and reference. Since her actual meaning—that she has eaten supper that evening—is an enrichment of the literal meaning, it counts as a case of literal use.[7]

If enrichment of meaning is to be seen as preserving literalness of use, then (40) must be treated as a case of literal use:

(40) [*Antony praising Brutus in Shakespeare's* Julius Caesar] This was a man!

Yet in classical rhetoric, (40) would be classified as a case of figurative use (more precisely, as a variety of synecdoche). Here again, intuitions are probably not clear enough to decide, so the decision should be made on theoretical grounds. However, we have argued that a notion of literalness has no role to play in a theory of language use. All utterances involve a process of meaning construction. This process is the same whether the outcome is an enriched, loosened, enriched-and-loosened, or literal interpretation. Yes, literalness can be defined, or at least characterised, in terms of a prototype, but, no, verbal understanding does not involve paying any attention to literal meaning, let alone to literal use. There are no theoretical grounds for sharpening our characterisation of literalness. On the other hand, as we will see, there may be social pressure to do so.

Similar problems arise with the notion of *what is said*. Given that a speaker produced some utterance *U* as an act of verbal communication, what is the proper completion of (41)?

(41) The speaker said that...

The idea that there is a theoretically clear and useful notion of *what is said* implies that there is one correct completion of (41), or a set of semantically equivalent completions,

[7] Note that, if it is to ground a clear notion of literalness, enrichment stands in need of exact definition. It cannot be defined in terms of entailment since, presumably, literalness of use is maintained under negation or embedding (e.g. in the antecedent of a conditional), whereas entailment relationships are not: *I have eaten* is entailed by *I have eaten supper tonight*, but *I have not eaten* is not entailed by *I have not eaten supper tonight*. Similarly with *If I have/haven't eaten (supper tonight), then P*).

that uniquely captures what is said by uttering *U*. This is, of course, compatible with recognising that there are many different ways of completing it that are pragmatically acceptable in different situations. For instance, the dots could be replaced by an exegesis, a summary or a sarcastic rendering of *U*. However, these would not fit the intended notion of *what is said*. Prototypical instances of the intended notion are easy to find: they are the same as the prototypical instances of *literal meaning*. For instance, what was said by the speaker of (37) is unproblematically rendered as (42):

(42) The anthropologist said that she has eaten human flesh.

Here, the speaker’s explicit meaning can be straightforwardly rendered by a transposition from direct to indirect quotation. However, this is not always so.

When Lisa, in (20b), utters "I have eaten" (with an explicature somewhat weaker than if she had uttered "I have eaten supper tonight"), what is she saying? Intuitions typically waver. *Saying* is commonly understood in an indirect-quotational sense, where what is said is properly rendered by an indirect quotation of the original utterance. It might thus be claimed that in uttering "I have eaten", Lisa is merely saying that she has eaten; but this would fail to capture the speaker's meaning. *Saying* can also be understood in a commitment sense, where what is said is what the speaker is committing herself to in producing an utterance. This sense is typically used when, precisely, the competence or the honesty of the speaker's commitment is being challenged. Suppose Alan replies to Lisa: "What you just said is false: I happen to know that you haven't eaten a thing since lunch!" By common-sense standards, Alan is not misusing the word "said". However, his response makes sense only if he takes Lisa to have said not just that she has eaten, but that she has eaten that very evening. Of course, Lisa might reply that she had so much lunch that she didn't feel like eating anything else that day. Although this might be seen as disingenuous, the explicature of her utterance is weak enough to leave room for reasonable doubt. If she had not eaten for days, then in uttering "I've eaten" in this situation, she would indisputably be saying something false.

The weaker the explicature of an utterance, the harder it is to paraphrase what the speaker said except in the indirect-quotational sense. Quoting the speaker's words, with or without transposition, is safe, but of limited use. Paraphrasing the speaker's meaning would be more useful were it not for the element of arbitrariness involved. This vacillation between a quotational and a commitment sense of saying is particularly obvious in the case of metaphor. On the one hand, we may feel that here the only safe sense of "saying" is the quotational one. When Mary utters "Holland is a picnic," we would all agree unhesitatingly that she is saying that Holland is a picnic. But this does not provide a truth-evaluable content which can be crisply and confidently paraphrased, and which Mary is clearly committing herself to. On the other hand, one can disagree with what is being said (in an everyday sense), without being able to paraphrase it. If Peter disagrees with the explicature of Mary's utterance, he may well tell her: "What you say is false!" Here, he would be expressing disagreement with Mary's explicature, however vague, rather than making the obvious point that Holland is a country and not a social event.

Speakers commit themselves, and they can be criticised for their commitments. Often, though, the exact character of their commitment can be contested. This happens quite commonly at home, in public life, and in court. Arguing about what was said—both its content and its truthfulness—is a social practice conducted within the framework of "folk-linguistics." *What is said* and *literal meaning* are folk-linguistic notions. Most people are more interested in the norms of linguistic communication than in its mechanisms. The apparent platitudes listed at the start of this article—as speakers, we expect what we say to be accepted as true, as hearers, we expect what is said to us to be true—are versions (one from the speaker's point of view, the other from the hearer's) of one of these folk-linguistic norms, a norm of truthfulness in what is being said.

As far as it goes, the norm of truthfulness is a rational requirement on verbal communication. It is generally invoked when the audience suspects that it is being violated, and it is very rare for a speaker accused of violating the norm to dispute its applicability. By contrast, disagreements about what was actually said are not rare at all. Appeals to literal meaning are typically made in the context of such disputes about what was said. It is often easier to agree on the literal meaning of an utterance, and on its literal truth or falsity, than on what the speaker meant, or what the hearer felt justified in understanding. Sometimes, a speaker can retreat behind the literal meaning of her utterance, which may have been true even if the utterance was misleading. At other times, hearers can point out that what was literally said was false, and the speaker may then argue that she was not intending to be taken so literally. Many such arguments are never settled. This is partly because it is a mistake to describe the speaker's commitment in terms of the folk-linguistic notion of *saying*.

The very idea that what a speaker says should invariably (except perhaps in the case of poetry) be either literal or else paraphrasable by a literal utterance is an illusion of folk-linguistics. Western folk-linguistics, at least, is committed to a code model of communication which entails that what is said should always be transparent or paraphrasable. Efforts to bring communicative practice into line this ideal have had some effect on language use. In forms of verbal interaction where speakers' commitments are particularly important from a social point of view (in science or law, for example), there is a demand that speech should generally be literal, and that occasional departures from literalness should be overt and blatant: occasional metaphors are acceptable but not the loose talk of ordinary exchanges. How well the demand is actually satisfied is another matter. Generally folk-linguistic theories about communicative practice have only peripheral effects on the natural processes of speech and comprehension, where so many of the sub-processes involved are automatic and impenetrable (see Levelt 1989).

It may have seemed reasonable to philosophers such as Paul Grice or David Lewis to use a reformulated norm of truthfulness as a cornerstone of their philosophy of language. Their reformulation did not go far enough. Both Grice and Lewis accepted that truthfulness based on the conventional meaning of utterances is expected. (For Grice, conventional meaning is just literal meaning; for Lewis, it is literal or figurative meaning, the latter being derived from the former.) This assumption played a central role in Lewis's explanation of how linguistic meaning could be conventional, and in Grice's account of how non-conventional meanings could be communicated.[8]

We agree that, at least in most cases, a hearer expects to be informed of something when he attends to an utterance. We agree with Grice that "false information is not an inferior kind of information; it just is not information" (Grice 1989: 371). So, yes, hearers expect to be provided with true information. But there is an infinite supply of true information which is not worth paying attention to. Actual expectations are of relevant information, which—because it is information—is (redundantly) true information. However, we have argued that there just is no expectation that the true information communicated should be literally or conventionally expressed.

Linguistically-encoded meaning is far too schematic and gappy to be capable of being true or false. It is just an input for further processing. Contrary to the standard view, this further processing does not consist in simply combining contextual reference with linguistic sense to determine a literal meaning. The fact that the speaker has produced this utterance with this linguistic meaning is expected to provide a relevant piece of evidence and a point of departure for inferring the speaker's meaning. The resulting explicatures and implicatures are in turn expected to provide worthwhile input for further processing: that is, to be relevant (and therefore true).

---

[8] Though her notion of a convention, and of the role of intention in communication, is opposed to those of Lewis or Grice, Ruth Millikan (1984) is similarly basing her philosophy of language on a version of the norm of truthfulness (see Origgi & Sperber forthcoming for a discussion of her approach).

References

Bach, K. (1994a) Conversational impliciture. *Mind and Language* 9: 124-62.

Bach, K. (1994b) Semantic slack: What is said, and more. In Tsohatzidis, S. (ed.). *Foundations of Speech-Act Theory: Philosophical and Linguistic Perspectives.* 267-91. London: Routledge.

Carston, R. (1988) Implicature, explicature and truth-theoretic semantics. In Kempson, R. (ed.). *Mental Representation: The Interface between Language and Reality*. 155-81. Cambridge: Cambridge University Press. Reprinted in Davis (1991). 33-51.

Carston, R. (1998) *Pragmatics and the Explicit-Implicit Distinction.* University College London PhD thesis.

Carston, R. (2000) Explicature and semantics. *UCL Working Papers in Linguistics*: 1-44. To appear in Davis, S. & B. Gillon (eds.). *Semantics: A Reader*. Oxford: Oxford University Press.

Carston, R. (forthcoming) *Thoughts and Utterances: The Pragmatics of Explicit Communication.* Oxford: Blackwell.

Davis, S. (ed.) (1991) *Pragmatics: A Reader*. Oxford: Oxford University Press.

Grice, H.P. (1989) *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Gross, S. (1998) *Essays on Linguistic Context-Sensitivity and Its Philosophical Significance*. Harvard University PhD dissertation.

Katz, J. (1990) *The metaphysics of Meaning*. Cambridge MA: MIT Press.

Keefe, R. & P. Smith (eds.) (1996) *Vagueness: A Reader.* Cambridge MA: MIT Press.

Levelt, W. (1989) *Speaking: From Intention to Articulation*. MIT Cambridge MA: Press.

Lewis, D. (1975) Languages and language. In Gunderson, K. (ed.). *Minnesota Studies in the Philosophy of Science*. Volume VII. 3-35. University of Minnesota Press.

Lewis, D. (1983) *Philosophical Papers*. Volume I. Oxford: Oxford University Press.

Lyons, J. (1977) *Semantics*. Volume 2. Cambridge: Cambridge University Press.

Millikan, R. (1984) *Language, Thought and Other Biological Categories*. Cambridge MA: MIT Press.

Neale, S. (1992) Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509-59.

Origgi, G. & D. Sperber (forthcoming) Evolution, communication and the proper function of language. To appear in Carruthers, P. & A. Chamberlain (eds.). *Evolution and the Human Mind: Language, Modularity and Social Cognition.* Cambridge: Cambridge University Press.

Recanati, F. (1989) The pragmatics of what is said. *Mind and Language* 4: 295-329. Reprinted in Davis (1991). 97-120.

Searle, J. (1979) *Expression and Meaning*. Cambridge: Cambridge University Press.

Sperber, D. (1994) Understanding verbal understanding. In Khalfa, J. (ed.). *What Is Intelligence?* 179-98. Cambridge: Cambridge University Press.

Sperber, D. & D. Wilson 1981 Irony and the use-mention distinction. In Cole, P. (ed.). *Radical Pragmatics.* 295-318. New York: Academic Press. Reprinted in Davis (1991). 550-63.

Sperber, D. & D. Wilson (1986a) *Relevance: Communication and Cognition.* Oxford: Blackwell. (Second edition 1995).

Sperber, D. & D. Wilson (1986b) Loose talk. *Proceedings of the Aristotelian Society* 1985-6: 153-71. Reprinted in Davis (1991). 540-49.

Sperber, D. & D. Wilson (1990) Rhetoric and relevance. In Wellbery, D. & J. Bender (eds.). *The Ends of Rhetoric: History, Theory, Practice.* 140-55. Stanford CA: Stanford University Press.

Sperber, D. & D. Wilson (1998a) The mapping between the mental and the public lexicon. In Carruthers, P. & J. Boucher (eds.). *Language and Thought: Interdisciplinary Themes.* 184-200. Cambridge: Cambridge University Press.

Sperber, D. & D. Wilson (1998b) Irony and relevance: A reply to Seto, Hamamoto and Yamanashi. In Carston, R. & S. Uchida (eds.). *Relevance Theory: Applications and Implications.* 283-93. Amsterdam: John Benjamins.

Stich, S. (1990) *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation.* Cambridge MA: MIT Press.

Travis, C. (1985) On what is strictly speaking true. *Canadian Journal of Philosophy* 15: 187-229.

Unger, P. (1975) *Ignorance: A Case for Scepticism*. Oxford: Clarendon Press.

Williamson, T. (1994) *Vagueness*. London: Routledge.

Wilson, D. (2000) Metarepresentation in linguistic communication. In Sperber, D. (ed.). *Metarepresentations: A Multidisciplinary Perspective*. 411-448. Oxford: Oxford University Press.

Wilson, D. & D. Sperber (1981) On Grice's theory of conversation. In Werth, P. (ed.). *Conversation and Discourse.* 155-78. London: Croom Helm.

Wilson, D. & D. Sperber (1992) On verbal irony. *Lingua* 87: 53-76.

Wilson, D. & D. Sperber (1993) Linguistic form and relevance. *Lingua* 90: 1-25.

Wilson, D. & D. Sperber (1998) Pragmatics and time. In Carston, R. & S. Uchida (eds.). *Relevance Theory: Applications and Implications.* 1-22. Amsterdam: John Benjamins.