

# BEYOND SPEAKER'S MEANING

Dan Sperber and Deirdre Wilson

CNRS/CEU, Budapest and UCL Linguistics/CSMN, Oslo

## Abstract

Our main aim in this paper is to show that constructing an adequate theory of communication involves going beyond Grice's notion of speaker's meaning. After considering some of the difficulties raised by Grice's three-clause definition of speaker's meaning, we argue that the characterisation of ostensive communication introduced in relevance theory can provide a conceptually unified explanation of a much wider range of communicative acts than Grice was concerned with, including cases of both 'showing that' and 'telling that', and with both determinate and indeterminate import.

**Key words:** communication, showing vs telling, paraphrasability, manifestness, ostension

## 1. Introduction

In *Relevance: Communication and Cognition* (Sperber and Wilson 1986; revised edition 1995) we put forward a number of novel ideas, several of which have been influential and others more controversial. However, there is one idea that we feel did not get the discussion it deserved. We proposed a characterisation of communication which, although inspired by Grice's definition of speaker's meaning, implied that speaker's meaning does not have the degree of unity or autonomy needed to make it the proper object of a philosophical definition or a scientific theory. Communication, on the other hand, or rather the kind of 'ostensive' communication that humans engage in, is such a proper object of inquiry. We argued that our account of communication does a better job of explaining how utterances are interpreted than a standard Gricean approach, and also makes good sense of our fuzzy intuitions about speaker's meaning without giving the notion an unduly important theoretical role. Here, we take up the issues again. In the first part of the paper, we discuss some difficulties with Grice's notion of speaker's meaning, and in the second part, we consider how to resolve them.

## 2. Difficulties with Grice's notion of speaker's meaning

### 2.1 *The continuum between meaning and showing*

Grice was aware of two particular problems with the notion of speaker's meaning. The first, linked to his sharp distinction between natural and non-natural meaning, arises when one tries to separate 'meaning that' from 'displaying direct evidence that'<sup>1</sup> in cases like the following (Grice 1989: 109):

- (a) Herod, showing Salome the head of St. John the Baptist, cannot, I think, be said to have meant that St. John the Baptist was dead.
- (b) Displaying a bandaged leg (in response to a squash invitation)

In (b) the displayer could mean (1) that he cannot play squash  
 Or (dubiously) (2) that he had a bad leg (the bandage might be fake)  
 But not (3) that the leg is bandaged.

In discussing case (a), Grice comments:

Herod intended to make Salome believe that St John the Baptist was dead and no doubt also intended Salome to recognise that he intended her to believe that St John the Baptist was dead... Yet I certainly do not think that we should want to say that we have here [a case] of meaning<sub>NN</sub> ... What we want to find is the difference between, for example, 'deliberately and openly letting someone know' and 'telling' ... (1989: 218).

Grice's solution was to add a third clause to his definition of utterer's meaning. In order to mean something by an utterance, the utterer must intend the addressee

- (1) to produce a particular response *r*
- (2) to think (recognise) that the utterer intends (1)
- (3) to fulfil (1) on the basis of his fulfilment of (2),

where (3) is understood as stipulating that the addressee's recognition of the utterer's intention in (1) must be "at least part of his reason for producing *r*, and not merely the cause

---

<sup>1</sup> Of course, if a reliable speaker both says that *P* and means it, this is evidence that *P*. However, it is indirect evidence (in the sense in which we want to make the direct/indirect distinction here) in that the content of the evidence depends on the interpretation of the communicator's meaning, and the force of the evidence depends on the communicator's trustworthiness.

of his producing *r*” (Grice 1989: 92). Despite some debate in the literature about whether this third clause was needed (Schiffer 1972; Vlach 1981; Recanati 1986; Bach 1987; Neale 1992; Wharton 2009), it remained central to Grice’s later discussions of meaning and his distinction between natural and non-natural meaning (Grice 1967/1989: 290–97, 349–59).

In characterising ostensive communication, we built on the first two clauses of Grice’s definition and dropped the third. This was not because we were willing to broaden the definition of utterer’s meaning – we agreed with Grice that talk of ‘meaning’ is awkward in certain cases – but because it seemed obvious that there is a continuum of cases between ‘meaning that’ (typically achieved by the use of language) and ‘displaying evidence that’ (in other words, showing), and we wanted our account of communication to cover both (Sperber and Wilson 1986: 46–54). Grice’s example of the bandaged leg suggests how the continuum of cases can be constructed. The communicator wants the addressee to come to believe that *P*. In pure cases of showing, as in (b3), what is being shown provides sufficient evidence for the addressee to believe that *P*, and the fact that the communicator intended him to believe that *P* does not even strengthen that evidence. In pure cases of meaning, as in an ordinary linguistic assertion, all the evidence that *P* is provided by the communicator’s giving overt evidence of her intention that the addressee should believe that *P* (and it is good evidence provided that the addressee trusts her competence and honesty in the matter). But of course, the evidence can come both from whatever is displayed (either shown or uttered) and from what the communicator’s communicative behaviour indicates of her intention, as in (b1) and (b2).

When a piece of evidence is shown to an addressee, it is typically interpreted in the light of the fact that it is being shown. In Grice’s example (b), the bandage on the leg may in itself be only weak evidence that the communicator cannot play squash: as he puts it, the bandage may be “fake”, and the condition it covers may be quite compatible with playing squash. However, the fact that it is being shown in answer to the question suggests that the condition it conceals makes playing squash impossible, or at least undesirable. Thus, (b1) is neither a pure case of meaning nor a pure case of producing direct evidence. Make the evidence stronger, say by showing a cast instead of a bandage, and the addressee will arrive at the intended conclusion mostly, if not wholly, on the ground of what is shown.

We described (b3) above as a pure case of showing. Could showing a bandaged leg ever ‘mean’ that the communicator has a bandaged leg (as some commentators have been willing to accept)? Suppose the showing was in response to the question, “Is your leg bandaged?”, how does this differ from simply answering “Yes”? One difference is that the addressee has

to trust the communicator in the ‘yes’ case and not in the showing case. Suppose, then, that the communicator pulls up her long skirt just enough to show what could be the bottom of a large bandage, giving weak, inconclusive evidence that her leg is bandaged. Since some trust is needed to accept the intended conclusion, would this now be a case of meaning? There is, of course, a continuum of positions to which the communicator could pull up her skirt, exposing a little more of the bandage each time, until the fact that the leg is bandaged is perceptually beyond doubt. At each point, less trust would be needed, exemplifying the continuum between meaning and displaying direct evidence.

Suppose Salome has never seen John the Baptist before, and is unable to recognise him. Then for her, seeing a severed head would not be compelling evidence that John is dead. When the head is shown to her by Herod, the evidence is stronger, because it is combined with recognition of his intention. This revised scenario seems to involve both direct and indirect evidence (i.e. both showing and meaning): Herod showed Salome that the person whose head he is displaying was dead, and meant that this person was John. Of course, Herod also overtly intended Salome to think that he was responsible for John’s death, and that he had had John killed to satisfy her wishes; since these were not wholly evidenced by John’s severed head, they must have been meant.

Perhaps Grice could have said that as long as recognition of the communicator’s intention plays a role – however small – in the addressee’s coming to the intended conclusion, the case is one of meaning. This seems to fit with his stipulation, in the third clause of his definition, that the audience’s recognition of the utterer’s intention should be “at least part” of their reason for producing the intended response. It would follow that any case of ‘showing that’ in which the evidence for the intended conclusion was less than decisive would have to be reclassified as a case of ‘meaning that’. But surely, if the part played by recognition of the utterer’s intention can vary from 100% to less than 1%, then many, if not most, cases of showing a piece of evidence seem to involve meaning, and the common sense understanding of meaning, and of the distinction between showing that and meaning that, is lost. A more sensible response would be to study the whole continuum – characterised by its two end points of pure meaning and pure showing – as such, and get rid of the third clause. However, this amounts either to extending the notion of speaker’s meaning way beyond what is intuitively recognisable as such, or to demoting it from its central theoretical role to a loosely descriptive use that may nonetheless be adequate when dealing with fairly standard cases of linguistic communication.

## 2.2 *The continuum between determinate and indeterminate 'meaning'*

A second difficulty Grice was aware of with the notion of speaker's meaning arises when one tries to complete a description of the form: "The speaker meant that \_\_\_\_". As Grice recognised, it is not uncommon for at least part of the intended meaning to be less than fully determinate, so that the best rendering of it may be an open disjunction of propositions, and hence not itself a proposition. As Grice put it (1989: 39-40),

Since, to calculate a conversational implicature is to calculate what has to be supposed in order to preserve the supposition that the Cooperative Principle is being observed, and since there may be various possible explanations, a list of which may be open, the conversational implicature in such cases will be an open disjunction of such specific explanations, and if the list of these is open, the implicatum will have just the kind of indeterminacy that many actual implicata do in fact seem to possess.

But this is tantamount to saying that there are some cases of speaker's meaning where "The speaker meant that \_\_\_\_" cannot be properly completed, not because the speaker failed to communicate a meaning, but because that meaning is not a proposition.<sup>2</sup>

We argued in *Relevance* that there is a continuum of cases from those where the communicator's meaning is a proposition, or can be paraphrased as such, to those where it is not paraphrasable at all. At one end of the continuum are utterances such as the railway official's reply to the passenger's inquiry below:

*Passenger:* What time is the next train to Oxford?

*Railway official:* 12.48.

Assuming that he has spoken in a neutral tone of voice and with an impersonal facial expression, his meaning could be paraphrased as the proposition that the next train to Oxford leaves at 12.48, and nothing more. Add an urgent tone of voice or a warning look, and although his assertion would remain the same, part of the intended import would be rather less determinate: he might be implicating, for instance, that the train is about to leave, that the

---

<sup>2</sup> In the 'Retrospective Epilogue', Grice described non-natural meanings as "conceptions or complexes which involve conceptions", and suggested that it would be legitimate to ask "how conceptions enter the picture and whether what enters the picture is the conceptions themselves or their justifiability" (Grice 1989: 350). This is very different from the picture normally presented in philosophy of language and linguistics.

seats are filling up fast, that the platform is further away than the passenger might have thought, that the passenger's estimated walking speed may not be enough to get her there on time, and so on. In that case, his meaning would be partly precise and partly vague.

With a hyperbole such as "I could kill for a glass of water", some of the speaker's words (e.g. "kill", "glass") are loosely used and no determinate proposition is asserted. Despite this element of indeterminacy, it is easy to see roughly what she is implicating: for instance, that she is very thirsty, that she has an urgent need or desire for water, and/or that getting hold of some water is a top priority for her. Although her meaning is less than fully determinate, identifying it is unlikely to give ordinary addressees much pause for thought.

With a poetic metaphor such as "Juliet is the sun", the speaker's meaning comes closer to the 'indeterminate' end of the continuum, and has the type of vagueness Grice saw as best rendered by an open disjunction of propositions. As Stanley Cavell comments (1965/1976: 78),

Romeo means that Juliet is the warmth of his world; that his day begins with her; that only in her nourishment can he grow. And his declaration suggests that the moon, which other lovers use as emblem of their love, is merely her reflected light, and dead in comparison, and so on. ... The 'and so on' which ends my example of paraphrase is significant. It registers what William Empson calls 'the pregnancy of metaphors', the burgeoning of meaning in them.

Vaguer still are non-verbal cases such as the following, taken from *Relevance* (Sperber and Wilson 1986/1995: 55):

Mary and Peter are newly arrived at the seaside. She opens the window overlooking the sea and sniffs appreciatively and ostensively. When Peter follows suit, there is no one particular good thing that comes to his attention: the air smells fresh, fresher than it did in town, it reminds him of their previous holidays, he can smell the sea, seaweed, ozone, fish; all sorts of pleasant things come to mind, and while, because her sniff was appreciative, he is reasonably safe in assuming that she must have intended him to notice at least some of them, he is unlikely to be able to pin down her intentions any further.

We went on to comment,

Is there any reason to assume that her intentions were more specific? Is there a plausible answer, in the form of an explicit linguistic paraphrase, to the question, what does she

mean? Could she have achieved the same communicative effect by speaking? Clearly not. (Sperber and Wilson 1986/1995: 55–6).

If asked what she intended to convey in this case, one of the best answers Mary could give is that she wanted to share an *impression* with Peter. Thus, at one end of the paraphrasability continuum are cases where the speaker's meaning is fully determinate, and at the other are those involving the communication of impressions, where the communicator's meaning cannot be paraphrased without loss.

### 2.3 *The two continua combined*

In raising these two issues, in arguing that there is a continuum between meaning and producing direct evidence and that paraphrasability is a matter of degree, we were not just being finicky: we were not making the trivial and boring point that there may be unclear, mixed or borderline cases along both dimensions. If that was the problem, a good theory of meaning and/or a good theory of showing could be used to arrive at theoretically-grounded decisions in unclear cases; and indeed, while awaiting the development of such good theories, one could ignore or idealize away fuzzy or borderline cases and investigate speaker's meaning by focusing on prototypical cases: that is, one could go on with philosophical business as usual. But our point was that to do this would be to idealise away essential features of communication, raising questions about the appositeness of the resulting theories.

Let us call the overtly intended cognitive effect of a communicative act its *intended import*. We want to argue that the two dimensions of intended import we are considering – meaning/showing and determinate/indeterminate – are orthogonal. As we have seen, completing the description “X meant that \_\_\_\_” with a proposition is sometimes unproblematic, sometimes impossible, and there is a continuum of cases in between. We will shortly demonstrate that the same point applies to “W showed that \_\_\_\_”. The two continua interact to yield a two dimensional-space, with intended imports occurring anywhere in this space. This is illustrated in Figure 1 and the examples immediately below:

**<Figure 1 about here>**

**Determinate meaning** (vicinity of 1). An example would be the railway official's reply "12.48" to the passenger's question about the time of the next train to Oxford, spoken in a neutral tone with an impersonal facial expression. This is a case of pure meaning, since all the evidence for the intended import comes from the speaker's intentions, and the meaning is determinate, since it is paraphrasable as a proposition. Most discussions of meaning in philosophy of language and linguistics focus exclusively on this type of case.

**Semi-determinate meaning** (vicinity of 2). An example would be the hyperbole "I could kill for a glass of water," where the intended import is vaguer than with the railway official's reply "12.48", but it is easy to see roughly what type of conclusions the addressee was intended to derive. This is a case of meaning, since all the evidence for the intended conclusions is indirect, but the meaning is less than fully determinate.

**Indeterminate meaning** (vicinity of 3). With a poetic metaphor such as "Juliet is the sun", the intended import is still vaguer, and is not paraphrasable as a proposition at all. This is again a case of meaning, since all the evidence for the intended import is indirect, but it is closer to the 'indeterminate' end of the paraphrasability continuum.

**Determinate meaning/showing** (vicinity of 4). When asked who is the tallest pupil in the class, the teacher points to an individual who at first sight is the tallest in the class (although some pupils might be absent) and says, "He is." She both means that the pupil she is pointing at is the tallest (since some of the evidence for the intended conclusion comes from her intentions), and displays direct evidence that he is the tallest. In both cases, the intended import is paraphrasable as a proposition.

**Semi-determinate meaning/showing** (vicinity of 5). On a tourist trip, Mary points to the view and says "What a view!" Here, the linguistic meaning of her utterance (combined with her tone of voice, facial expression etc.) indicates roughly what type of conclusions she expects the addressee to derive, but does not pin them down precisely, so the utterance falls towards the middle of the 'determinate/indeterminate' continuum. Moreover, the evidence for the intended conclusions comes both from Mary's intentions and from what she has pointed out, so the utterance also falls towards the middle of the 'meaning/showing' continuum.

**Indeterminate meaning/showing** (vicinity of 6). On a tourist trip, Mary points to the view and says "Wow!" This time, the linguistic meaning of her utterance (to the extent that it has one) gives no more than a rough indication of the type of conclusions the addressee is being



encouraged to derive, and the intended import is not paraphrasable as a proposition at all. In Grice's terms, what Mary communicates is an open disjunction of propositions; in our terms (to be discussed further below), what she communicates is an impression.

**Determinate showing** (vicinity of 7). When asked for the time, Mary points to a clock showing the time as five o'clock. Here, the intended import is as determinate as if she had said "It's five o'clock". However, the case is one of showing rather than meaning, since all the evidence for the intended conclusion comes from the clock itself, rather than from the fact that it has been pointed out.

**Semi-determinate showing** (vicinity of 8). Peter and Mary are out for a walk when she points to menacing clouds on the horizon. Here it is easy to see roughly what she intends to convey – that it may rain soon, that they should reassess their plans and maybe think about curtailing their walk – but the intended import is less than fully determinate.

**Indeterminate showing** (vicinity of 9). An example might be showing pictures of one's children. Here, there is no proposition that would complete the description "The communicator showed that \_\_\_", and the intended import cannot be rendered as a proposition at all.

Notice that it is possible to mean and show the same thing, as when the teacher, asked who is the tallest pupil in the class, points to the tallest individual in the room and says "He is". This allows us to handle a type of example that seems to be incompatible with the third clause of Grice's definition of meaning, and led him to contemplate dropping this clause entirely. When the communicator is producing a logical argument, she typically intends her audience to accept the conclusion of this argument not on her authority, but because it follows from the premises (this type of case is insightfully discussed by Schiffer 1972):

*Conclusion of argument:  $p, q, \text{ therefore } r$  (from already stated premises):* While  $U$ [*tterer*] intends that  $A$ [*addressee*] should think that  $r$ , he does not expect (and so intend)  $A$  to reach a belief that  $r$  on the basis of  $U$ 's intention that he should reach it. The premises, not trust in  $U$ , are supposed to do the work. (Grice 1989: 107).

Since the third clause was crucial to maintaining Grice's distinction between 'meaning that' and 'displaying direct evidence that', he was reluctant to drop it. We would analyse this type of example as a case of determinate meaning/showing. The speaker provides both direct and

indirect evidence that the conclusion follows from the premises: that is, she both means it and shows it.

There is another type of case that Grice did not discuss, but that does raise a serious problem for his distinction. Many utterances contain deictic elements whose function is not just to specify a referent but also to specify conceptual content (the referent being a token that contributes to the interpretation a type to which it belongs). Compare, for instance:

To open a champagne bottle, you can do this (*demonstrates how to open a champagne bottle by opening one*)

To open a champagne bottle, you should do this (*demonstrates how to open a champagne bottle by opening one*)

The first communicative act (comprising both the utterance and the demonstration) does not satisfy the third clause of Grice's definition of speaker's meaning, since the demonstration provides sufficient evidence of the fact that a bottle of champagne *can* be opened in the way demonstrated. By contrast, the second communicative act (where the only difference is that 'can' has been replaced by 'should') is a perfect case of Gricean speaker's meaning, since the demonstration is not sufficient evidence for the normative claim. But of course the two acts are very similar in their communicative import, and should be analysed in very similar ways.

This example also illustrates the fact that deixis which helps to specify conceptual content is quite commonly a source of indeterminacy. In a communicative act of this type, the demonstrative behaviour is merely indicative of what the speaker intends to convey. Some of its features should be replicated in future performances, and others not. As in most 'how-to' demonstrations, the movements are individually highlighted in a way that is useful to the demonstration, but not to the opening of a champagne bottle. This highlighting is not to be replicated in future performances. Moreover, the demonstrator has her own idiosyncrasies – she may be left-handed, for instance – that need not be replicated either. To comprehend such a communicative act involves inferring the type of action of which the demonstration was an ostensive, and hence untypical, token, and understanding that the 'this' denotes the type. Quite commonly, such demonstrations are used because no perspicuous verbal description is available – in which case, the content communicated is not paraphrasable.

This example shows that not only meaning and showing but also determinate and indeterminate aspects of the intended import can coexist in a communicative act. The co-

occurrence of precise and vague import is also common in ordinary verbal communication, as when the railway official's reply "12.48" is accompanied by a warning tone or look.

In the next section, we consider the cognitive background against which communication takes place, reviewing several distinctions among types of mental state that are relevant to our discussion. In the following two sections, we introduce a theoretical notion, manifestness, which helps to clarify the relations among these various types of mental state. Finally, we will use the notion of manifestness to characterise ostensive communication, and apply the resulting framework to some of the examples discussed in sections 2 and 3.

To keep the discussion brief, we will consider only declarative (as opposed to directive) acts, for instance, saying "It's five o'clock", or pointing to a clock showing the time as five o'clock. In *Relevance* and elsewhere, we have suggested how to extend the analysis to other kinds of illocutionary act (Sperber and Wilson 1986/1995, chap. 4, section 7; Wilson and Sperber 1988/2012).

### 3. Types of mental state: beliefs and impressions

In his earliest work on meaning, Grice assumed that declarative acts were intended to induce a belief in the audience (e.g. the addressee of the utterance "It's five o'clock" would be intended to form the belief that it was five o'clock). As he later recognised, however, reminders and recapitulations present problems for this approach (Grice 1967/1989: 106–7):

*Reminding:*        Q: "Let me see, what was that girl's name?"  
                          A: "Rose" (or produces a rose).

The questioner is here already presumed to believe that the girl's name is Rose (at least in a dispositional sense); it has just slipped his mind. The intended effect seems to be that A should have it in mind that her name is Rose.

*Review of facts:* Both speaker and hearer are supposed already to believe that  $p$  ( $q$ , and so forth). The intended effect again seems to be that A (and perhaps  $U$  also) should have "the facts" in mind (altogether).

In response, Grice (1967/1989: 109) suggested that declarative acts might be intended to induce not just a belief but an 'activated belief' (in his terms, a belief that the addressee not only has, but "has in mind"). An addressee might fall short of having an activated belief in one of three ways. He might:

- (1) neither believe that  $p$  nor have it in mind that  $p$
- (2) believe that  $p$  but not have it in mind that  $p$
- (3) not believe that  $p$  but have it in mind that  $p$

Ordinary assertions, reminders and recapitulations could then be seen as addressing different types of shortfall, inducing activated beliefs by different routes.

Grice's suggestion raises a more general question about distinctions among types of belief (roughly captured by contrasts such as 'activated/latent', 'occurrent/dispositional', 'explicit/implicit'), and the extent to which declarative acts are intended to induce some specific type of belief in the audience. We will argue that the typologies of belief-states common in philosophical psychology may not be adequate to answer this question.

### 3.1 *Occurrent, dispositional and implicit beliefs*

Beliefs are commonly seen as playing a central causal role in the explanation of thought and behaviour. In the kind of accounts we are concerned with here, a belief is a representation that has to be occurrent or activated in the mind in order to play such a causal role. Occurrent or activated beliefs contrast with inactive or dispositional beliefs, which are also understood as being 'in the mind' (in a different sense from Grice's), although not immediately available for use as premises in theoretical or practical inferences.

A somewhat psychologically richer way of describing this distinction might be to say that activated beliefs are in working memory, whereas latent beliefs are in long-term memory and have to be retrieved in order to play a causal role. Or, to borrow a metaphor from Robert Audi (1994: 420),

What is dispositionally as opposed to occurrently believed is analogous to what is in a computer's memory but not on its screen: the former need only be brought to the screen by scrolling – a simple retrieval process – in order to be used, whereas the latter is before one's eyes. Compare a dispositionally believed proposition's needing to be "called in," as in answering a request to be reminded of what one said last week, with an occurrently believed proposition's being focally in mind, roughly in the sense that one attends to it, as where one has just formulated it to offer as one's thesis.

Activated and latent (or occurrent and dispositional) beliefs are seen as representations 'in the mind'. Both are also described as 'explicit', and contrast with contents that an agent may be

said to believe even though they are not represented in her mind; these are sometimes called ‘implicit’, or ‘tacit’, beliefs.

Here is how Eric Schwitzgebel (2006) presents the distinction between explicit and implicit beliefs in the *Stanford Encyclopedia of Philosophy*:

One believes *P explicitly* if a representation with that content is actually present in the mind in the right sort of way – for example, if a sentence with that content is inscribed in the “belief box”... One believes *P implicitly* (or *tacitly*) if one believes *P*, but the mind does not possess, in a belief-like way, a representation with that content.<sup>3</sup>

This raises the question of how to distinguish implicit beliefs from contents that are not themselves believed, although they follow logically from one’s beliefs. Schwitzgebel suggests that there may be no clear cut-off point between the two:

Perhaps all that's required to implicitly believe something is that the relevant content be swiftly derivable from something one explicitly believes ... Thus, in the planets case, we may say that you believe explicitly that the number of planets is 9 and only implicitly that the number of planets is less than 10, less than 11, etc. Of course, if swift derivability is the criterion, then although there may be a sharp line between explicit and implicit beliefs (depending on whether the representation is stored or not), there will not be a sharp line between what one believes implicitly and what, though derivable from one's beliefs, one does not actually believe, since swiftness is a matter of degree.

In keeping with this suggestion, Robert Audi (1994: 419) argues that what are generally called implicit beliefs are better viewed not as beliefs at all, but as *dispositions to believe*:

Do you believe that this sentence has more than two words? And do you believe that 98.124 is larger than 98? It would be natural to answer affirmatively. And surely, for most readers considering these questions, that would be answering truly. [...] Antecedent belief may be the readiest explanation of our spontaneous answers, but it is not the best explanation. I contend that, here, what may seem to be antecedently held but as yet

---

<sup>3</sup> Schwitzgebel notes that some philosophers use the term ‘dispositional’ for what he is calling implicit beliefs. However, he reserves the term for latent (as opposed to occurrent) beliefs, and we will follow him on this.

unarticulated dispositional beliefs are really something quite different: dispositions to believe.

Integrating these ‘dispositions to believe’ into his computer screen metaphor, Audi writes:

By contrast with both of these cases of actual belief [i.e., occurrent and dispositional beliefs], propositions we are only disposed to believe are more like those a computer would display only upon doing a calculation, say addition: the raw materials, which often include inferential principles, are present, but the proposition is not yet in the memory bank or on the screen. The suggested difference between a dispositional belief and a disposition to believe is in part that between accessibility of a proposition by a retrieval process that draws on memory and its accessibility only through a belief formation process.

As a result of these considerations, one might divide beliefs into three categories, as in Figure 2:

**<Figure 2 about here>**

The problem with this proposal is that what actually exists, at least in the case of dispositional beliefs and dispositions to believe, are not distinct categories but a continuum of cases, Arguably, we do actually have activated representations in working memory. However, there is a long-standing consensus in the psychology of memory that what we have in long-term memory is not a repertoire of representations that we can simply move to working memory (or scroll down to), but traces or bits of information from which actual representations are reconstructed for use in working memory. This is not to deny that there are also likely to be some full-fledged representations which can simply be activated (as when you remember the Pythagoras theorem you learned by heart at school). But when you remember facts about the last departmental meeting – say, that John spoke after Jean, and that she seemed irritated – you are not simply moving these representations from long-term to short-term memory: the chances are that they were not stored there as distinct representations in the first place. What you have in long-term memory is information from which these representations can be accurately constructed – as opposed to just being pulled out. With some pieces of memorized information, retrieval – a misleading term – involves more inferential reconstruction than

with others. But the point is that, since retrieval from long-term memory typically involves some inference, it is not possible to distinguish implicit beliefs from latent/dispositional beliefs on the ground that one is derived via inference and the other is not.

Here is a continuum between dispositional and implicit beliefs in the spirit of Audi's argument that we merely have a disposition to believe that 98.124 is larger than 98. If the only requirement for a belief of yours to be dispositional is that you have memorized this information at some point in the past and are now remembering it, then your belief that 9 is larger than 8 is surely dispositional: the sequence of numbers from 0 to some small number is permanently represented in your mind.<sup>4</sup> Your belief that 99 is larger than 98 might also be dispositional, since you have probably also entertained it (in some form), for instance when counting to 100. What about your belief that 7899 is larger than 7898? Well... For some  $n$ , your belief that  $n$  is larger than  $n-1$  has been previously represented in your mind, whereas your belief that  $n+1$  is larger than  $n$  has not (and of course you don't know the value of  $n$ ). However, there is no interesting epistemic or psychological difference between beliefs about successive numbers lower than  $n$  and beliefs about successive numbers higher than  $n$ . Moreover, reaction time studies show that such beliefs are constructed on the basis of a mental number line, in a way that goes quite against the idea that previously held beliefs are merely reactivated (Dehaene 1997). These studies provide evidence that, for any two numbers  $m$  and  $n$ , it takes more time to answer the question, "Is  $m$  larger than  $n$ ?" when  $n$  and  $m$  are closer than when they are further apart. Thus, it takes more time to answer 'Yes' to the question "Is 29 larger than 28?" than to the question "Is 69 larger than 28?" Yet when  $n$  and  $m$  are close, and even more so when they are adjacent, it is much more likely that the proposition  $m$  is larger than  $n$  has already been entertained than when  $n$  and  $m$  are distant. So if activating a belief was a matter of retrieving past representations, the answer should be faster for adjacent numbers.

With numerical examples of this type, the propositions we are disposed to believe follow logically from what we explicitly believe. But we are also disposed to believe propositions that follow non-demonstratively from our mentally represented beliefs. So, for instance, we implicitly believe (or are disposed to believe, and in any case would assent to the claim) that the weather in New York will be warmer next July than next January, that more people were born in 1992 than in 1932, that Helsinki is east of Naples, and so on.

---

<sup>4</sup> This is not the same as having the relation between any two successive numbers in that sequence itself stored as a distinct representation, but we will ignore this for the time being.

How readily we assent to some statement that does not express an activated belief of ours depends on two factors, one epistemic, and the other cognitive. On the epistemic side, we don't simply believe or not believe a proposition: we believe it more or less strongly. The less strongly we believe it, the less willing we may be to assent to it when expressed, and the less appropriate it is to describe our attitude to it as one of belief. The point being that here, too, there is a continuum, between propositions we believe and propositions we neither believe nor disbelieve, with no cut-off point or even a bimodal distribution of instances. On the cognitive-processing side, a given dispositional or implicit belief of ours may be more or less salient, more or less easy to reconstruct or infer. The salience of a belief is not just a function of its epistemic strength: some logical entailments of what we strongly believe may not be salient at all, whereas a merely probable implication of what we weakly believe may be highly salient. For instance, when someone tells us that her sister Jane is in town, the implication that Jane or the Pope is in town is unlikely to become very salient, although it follows logically from what we have been told (and may quite strongly believe). By contrast, when someone tells us that a neighbour of theirs may have Alzheimer's, the probable implication that the neighbour is old is likely to be quite salient, although our evidence for this is not conclusive. So dispositional and implicit beliefs are on a continuum with propositions we do not believe in any sense (activated, dispositional or implicit), either because they are too hard for us to infer, or because we are not disposed to give them enough credence.

There are also propositions that we are disposed to believe although they don't follow either logically or non-demonstratively from what we already believe, but follow instead from what we already believe together with what we perceive. Consider Audi's example: "Do you believe that this sentence has more than two words?" It is conceivable (just) that on reading this sentence, you formed the mentally-represented belief that you were reading the sentence, 'Do you believe that this sentence has more than two words?' and are able to infer from that belief, together with your ability to inspect it in memory, that the sentence has more than two words. What about, "Do you believe that there are more than 50 words in this paragraph?" You will probably answer 'Yes', but the way you arrive at this answer will be by looking at the paragraph rather than consulting your memory. Here again, there is a continuum of cases between those where you have a disposition to believe immediately, and those where, in order to answer the question, you have to attend perceptually to more than you were attending to already. In these latter cases, the disposition to believe is not (so to speak) wholly inside you, but also involves the environment. However, the fact that the



environment is involved does not stop it being a disposition. This environment-dependent disposition to believe may in fact be stronger than a purely internal disposition to believe.

With all this in mind, let's return to Grice's suggestion that a declarative act might be intended to induce not just a belief but an 'activated belief'. In the case of reminders, this suggestion seems quite plausible; but how far does it generalise? Consider the following dialogue:

*(John is offering drinks to his guests; some have already taken whisky, vodka, cognac or orange juice)*

*John (to Rita):* Do you want some whisky?

*Rita:* I don't drink alcohol.

Rita is explicitly communicating that she doesn't drink alcohol and implicitly communicating several further propositions:

- (a) She doesn't want whisky
- (b) Her reason for refusing his offer of whisky is that she doesn't drink alcohol
- (c) She doesn't want cognac
- (d) She doesn't want vodka
- (e) She might accept orange juice

But does Rita intend John to entertain all these implicatures as activated beliefs? Certainly, she intends him to form the activated belief that she doesn't want whisky (implicature (a)). Of the other implicatures, though, she would typically intend him to activate only those that may turn out to be relevant to his actions. Suppose that, instead of forming the activated belief that her reason for refusing his offer is that she doesn't drink alcohol (implicature (b)), he merely holds it dispositionally, to await activation should the need arise. Even so, it would be quite wrong to say that communication has failed: a belief can be communicated without being activated in the addressee. In appropriate circumstances – for instance, if it crosses his mind to offer her vodka instead of whisky – he may form the activated belief that she doesn't want vodka (implicature (d)); otherwise, the fact that he is disposed to form this belief as a result of her utterance should be enough for communication to succeed.

This description of the different kinds of credal disposition an utterance may cause in an audience is sensitive to distinctions Grice does not envisage; however, it is still not fine-grained enough. In many cases of verbal or non-verbal communication, what the

communicator wants to do is not to induce a specific belief or set of beliefs in the audience, but to cause what might be roughly described as an impression, giving rise to a range of non-paraphrasable effects. Grice's suggestion that what is conveyed in this type of case might be analysed as an "open disjunction" of propositions is not really helpful; we will try to improve on it here.

### 3.2 Impressions

What is an impression? In section 2, we used the example of Mary, newly arrived on holiday, sniffing appreciatively and ostensibly at the fresh seaside air in order to share an impression with Peter. Here are two more examples:

Robert, working at his desk, is wondering whether to take a break and go for a walk. He gets up, opens the window: the sky is grey; the air is chilly; clouds, some of them rather dark, are moving fast. The impression he forms of the conditions outside make him change his mind. He will stay at home.

John has told Julia – who believed him – that the artist is a pretentious and rather conventional painter. However, he has to go to the exhibition, and he begs her to come too. As she walks through the gallery, she is pleasantly surprised by several of the paintings. Although she couldn't have pinpointed what she likes about them, she finds them arresting and somehow insightful. The impression she forms makes her change her mind. What John has told her is false. The painter, she comes to believe, is original and talented.

As these examples show, the formation of an impression, just like the formation of a belief, can bring about a theoretical or practical change of mind.

Bringing about a change of mind in one's audience is a typical goal of communication: indeed, with a suitably extended understanding of 'change of mind', it is the goal of any act of communication. The goals of communication, then, can sometimes be achieved by conveying an impression. For instance, Robert says to Susan, "Let's go for a walk." She might answer by opening the window and showing him the poor weather outside. She would thereby be causing him to form an impression about the weather, and doing so ostensibly (i.e. in such a way as to let him know that she is trying to communicate with him). Robert could then infer from his impression that it is not a good idea to go for a walk, and moreover that she intended him to form that impression and come to that conclusion, both of which

they now share. Or as they walk through the exhibition, Julia might say to John: “I don’t know how to put it... not what I expected... these paintings have something..., I’m sure you too must... Thanks for making me come!” As a result of her utterance, John might form an impression of Julia’s impression, which in turn might help him share it, or at least revise his first impression in the direction of hers. In many circumstances, this might be quite enough for Julia’s communication to succeed.

But what exactly is an impression? Does this common-sense notion pick out some mental state that we should be able to describe in cognitive terms, or is it irrelevant to scientific psychology? In any case, what is the process commonly described as forming an impression, and how does it achieve its cognitive impact? How can an impression be what a communicator wants to convey? How is it triggered and exploited in communication?

#### **4. Manifestness**

What occurrent, dispositional and implicit ‘beliefs’ have in common is that there is some proposition that you are likely to some positive degree to entertain and accept as true. Following our proposal in *Relevance*, we will say that this proposition is *manifest* to you. Manifestness depends on two factors mentioned in section 3.1 above: strength of belief and salience.<sup>5</sup> These factors are quite different – one is epistemic and the other cognitive – and for some purposes it would be unsound to lump them together. However, we need to consider their joint effect in order to explain or predict the causal role of a piece of information in the mental processes of an individual. The greater the degree of manifestness (i.e. the resultant or vector sum of these two factors, epistemic strength and salience) of some piece of information to an individual, the greater the causal role of that information in the individual’s thought and behaviour.

Here, then, is a definition of manifestness that differs marginally in formulation, though not in import, from the one given in *Relevance* (Sperber and Wilson 1986/1995: 38–41), followed by some clarificatory comments:

##### *Manifestness*

A proposition is manifest to an individual at a given time to the extent that he is likely to some positive degree to entertain it and accept it as true.

##### **Comments:**

---

<sup>5</sup> In *Relevance*, we used the term ‘accessibility’ rather than ‘salience’ to refer to the same property.

- In *Relevance*, we talked not of propositions but of assumptions as being manifest. Either term will do, and so would ‘pieces of information’. What we are talking about are things that can be true or false and that, when they are true, are facts.
- The time referred to is not a time point but a time span within which an inferential process may take place. Since such processes can be more or less extended in time, the time span may vary from a fraction of a second to a much longer period in which a protracted inference process such as a scientific discovery takes place. What makes a proposition manifest in such a time span is not only the mind-brain state of the individual during that time, but also the environment of the individual and the information it provides him with via perception and communication. Of course, if we are talking about what is manifest to Robert at the moment he opens the windows, or what is manifest to Julia as she walks through the exhibition, or what was manifest to Eratosthenes while he was calculating the circumference of the earth, the situations are quite different. Depending on the time span, what is manifest to the individual may involve information provided by the environment, or communicated by others, to a lesser or greater degree. Another way of putting this is that the mental processes involved may be more or less extended not only in time, but also in physical and social space, and include processes characteristic of what is described in the literature as embodied, situated, or distributed cognition. We see this not as a problem with our definition but as an advantage. Here, though, we will be concerned with the kind of short time span and organism–environment interactions involved in understanding an utterance.
- In *Relevance*, we described the two factors that contribute to the manifestness of a proposition: its salience (‘accessibility’) and the degree to which it is accepted as true. These two factors affect the probability that a proposition will influence an individual’s beliefs or decisions: the higher the probability that it will be accessed, the higher, *ceteris paribus*, the probability that it will have some influence, and the higher the degree to which it is accepted as true, the stronger that influence. However, manifestness is this *ceteris paribus* probability of influence, rather than the factors that contribute to it. The same point can be made on the basis of the Cognitive Principle of Relevance (‘Human cognition tends to be geared to the maximisation of relevance’): Given a belief which has a cognitive effect and which is therefore relevant in a categorical sense, its relevance will be comparatively greater 1) to the extent that its

processing is less costly because it is more salient, and 2) to the extent that its epistemic strength is greater. If a proposition is relevant at all, then the greater its manifestness, the greater its relevance.

In *Relevance*, based on this notion of manifestness, we introduced several further notions:

A *cognitive environment* of an individual at a time is a set of assumptions/ propositions that are manifest to that individual at that time. A cognitive environment can be shared between two or more individuals if it is a cognitive environment of each of them. Among the propositions manifest in a *shared cognitive environment*, some may enumerate or identify the people who share that environment. In that case, this shared environment is also a *mutual cognitive environment*, and all the propositions in it are mutually manifest. In *Relevance*, we spend some time showing that the notion of *mutual manifestness* is more realistic, more psychologically relevant, and at least as cogent as the notions of mutual knowledge, common knowledge or common ground (Lewis 1969; Clark and Marshall 1981; Sperber and Wilson 1986/1995: chapter 1, sections 3, 8; 1990; Stalnaker 2002) used in particular to explicate the Grice–Strawson insight that communicative intentions are ‘overt’. We will not discuss these uses of the notion of manifestness here.

We also used the notion of manifestness to redefine the content of communicative intentions so as to provide an explicit and unitary account of cases involving the communication of single propositions, on the one hand, and of what Grice describes as “open disjunctions”, on the other. This is the use of ‘manifestness’ that we will now elaborate and discuss.

We will begin our account in the next section by reassessing the types of mental state discussed in section 3 in the light of this characterisation of manifestness. We will argue, first, that the rather artificial categories of occurrent/ dispositional/ implicit beliefs are on a continuum of degrees of manifestness, and second, that we can give a better account of what it is to have an impression using the notion of manifestness.

## **5. Manifestness, beliefs and impressions**

### *5.1 Beliefs and manifestness*

Occurrent beliefs are not all equally manifest. In the first place, they are not all equally salient. Suppose you are asked to recommend your two favourite restaurants. You activate a number of beliefs about each, but some of them stand out more than others: for instance, your

belief that the desserts at Pierre's are exceptional may be the most salient of all your beliefs about either restaurant, whereas your belief that it is hard to park at La Cantina, although activated, may not be much attended. In the second place, occurrent beliefs may be epistemically stronger or weaker: you might think of both restaurants as having an excellent wine list, but be more convinced of this in the case of La Cantina than of Pierre's. *Ceteris paribus*, the activated beliefs that are more salient, and those that are more firmly held, are more likely to inform your conclusions.

At a given moment, there may be a genuine functional difference between occurrent beliefs in working memory and dispositional beliefs in long-term memory. However, since the contents of working memory are constantly changing, with some elements being added and others dropping out, a belief that is about to be recovered from long term memory may well play a greater role in a given inference than a belief that is still in working memory (though not for long) because of the role it played in a preceding inference. If we look at the mental status of beliefs dynamically (as we should), it should be clear that an occurrent belief is not necessarily more manifest than a dispositional one. Dispositional beliefs of course vary along the dimensions of salience and strength, and hence of manifestness. Some of our dispositional beliefs may well be less manifest – indeed, much less manifest – than implicit beliefs we have never entertained before but for which there is strong and salient evidence, either in memory or in the environment. For instance, you may be able to answer the question, “Would the children prefer Pierre's or La Cantina?” immediately and with confidence, even though you had never previously entertained the answer you now give; by contrast, it may take you some time to answer the question, “What is the name of the chef at La Cantina?” and your answer may not be entirely confident, even though you have heard the name before (*hum, I remember commenting that it was not an Italian name but a Hungarian one, a famous name actually, Kadar? Molnar? Lukács? Yes, Lukács, like the philosopher, I think*).

To sum up, the division of beliefs or potential beliefs into three categories is too coarse. A gradient in terms of manifestness is more helpful.

## 5.2 Impressions and manifestness

Using the notion of manifestness, we can also give a more precise account of impressions. When Robert, intending to go for a walk, opens the windows to see what the weather is like and alters his plans, what happens to, and in, his mind? We might be tempted to say that, on the basis of his perceptions, he has formed new beliefs and used them as premises in a

practical inference. Which new beliefs? Well, maybe the belief that the sky is grey and the air is quite cold, that it is therefore likely to rain, and that the weather is not right for taking a walk. Many combinations of similar and related beliefs would indeed warrant the practical conclusion he has arrived at. Although we cannot know exactly which beliefs Robert formed and used as premises, we might assume that only such a determinate set of beliefs could have caused Robert to change his mind. It might suit our theoretical agenda to think that there are facts of the matter, and that Robert knows them, or may have known them, however fleetingly, at the time.

But we have all been in Robert's position, and we should envisage the possibility that he might have come to his decision without ever being aware of clear and distinct premises, or of deriving his decision from those premises. Here is an alternative description. When Robert opened the window, an array of propositions became manifest or more manifest to him, in the sense characterised above: they became more likely to be attended to, and more likely to be taken as true, than they had been before, and were therefore more likely to influence his decision. He may have been aware of this increase in the manifestness of an array of propositions, and of their general drift, without entertaining all of them, and maybe even without entertaining any of them as a distinct proposition, except for the practical conclusion that he would not go for a walk.

Not all inferences involve step by logical step derivations of explicit conclusions from explicit premises. Arguably, the vast majority of inferences made by humans and other animals do not involve such derivations. What happened in Robert's brain when he opened the window might be better described as changes in patterns of activation, none of which would properly speaking amount to the fixation of a distinct credal representation, but the totality of which would correspond to the formation of an impression. These changes would then jointly inhibit what may have been a distinct volitional representation, his desire to take a walk. Thus, rather than a step-by-step derivation of an explicit conclusion from explicit premises, the inferential process might have consisted in John's impression of the weather undermining his desire to go for a walk. More generally, many (if not all) inferences can be described not as more or less standard logical derivations but as competitions between alternative conclusions (it will rain/it won't rain, let's go for a walk/let's not, and so on). The winner of such competitions is determined by activation or inhibition caused by brain states that represent information in all kind of ways (from consciously entertained propositions to unconscious weightings of features – where 'represent' is broadly understood as meaning *fulfil the function of making some information available for processing*). If the mental

mechanisms which decide the outcome of such competitions tend to favour warranted conclusions, then although the process is quite different from a sequence of good old syllogisms, it would still be genuinely inferential. We are not arguing for this view of human inference here, but merely arguing against tying our understanding of the role of inference in communication to an old, narrow and questionable view of what such inference must be like.

We are suggesting, then, that an impression is a change in the manifestness of an array of propositions which all bear on our understanding the same phenomenon, answering the same question, or deciding on the same issue. This is not intended as an analysis of the ordinary use of the word ‘impression’. If someone told us that this account of impressions does not capture the ordinary usage of the term, we would say that we are proposing to use ‘impression’ as a technical term, to denote a psychologically relevant category, whether or not this category is recognised as such in common-sense or philosophical psychology. In fact, though, we believe our technical use corresponds fairly closely to – and, if anything, sharpens – the typically vague common-sense meaning of ‘impression.’ It may be worth briefly showing this by considering a possible objection. ‘To have the impression that’ can be construed as expressing a propositional attitude that takes a proposition as its complement. For instance, Robert might say “I had the impression that it would rain”. Does this use of ‘impression’ differ in meaning from the one we were trying to capture above? One might be tempted to say that ‘to have the impression that’ denotes a weak credal state, not quite a belief that *P*, but a belief that *probably P*. However, this is demonstrably wrong. Suppose you enter a classroom and say, “I have the impression that there are more than fifty and fewer than a hundred people in this room”. This is an appropriate use of the phrase, and corresponds to a situation where you have formed an impression in the sense we described earlier, from which a conclusion follows. That is, an array of propositions have become manifest to you, and although you are not aware of them individually, this overall change in your cognitive environment warrants the inference that there are probably more than fifty and fewer than a hundred people in the room. Now suppose you were to say instead, “I have the impression that the number of people in the room is not a multiple of 11,” drawing on your knowledge of the fact that the chance of a random integer being a multiple of 11 is one in eleven. In this case, your use of the phrase would not be appropriate, even though you would be expressing the attitude of taking the proposition embedded in your utterance to be probably true. Thus, ‘impression’ does not simply pick out a weak credal attitude; it picks out a certain type of vague information basis for such an attitude.



We now have all we need to address Grice's worries about the kind of effects a declarative act is intended to induce in the audience, and to provide a framework in which the full range of possible effects – including prototypical cases of speaker's meaning – can be treated in a conceptually unified way.

## 6. Manifestness and communicative intentions

In *Relevance*, we began by proposing an informal and incomplete definition of the two intentions involved in ostensive communication (corresponding to the first two clauses of Grice's definition of meaning) (Sperber and Wilson 1986/1995: 29):

*Informative intention*: to inform the audience of something;

*Communicative intention*: to inform the audience of one's informative intention.

Using the notion of manifestness, we then gave a more precise and fuller version which we reformulate slightly here (Sperber and Wilson 1986/1995: chapter 1, section 10):

*Informative intention*: to make manifest or more manifest to the audience an array of propositions **I**.

*Communicative intention*: to make it mutually manifest to audience and communicator that the communicator has this informative intention.

We won't discuss here the role of mutual manifestness in the communicative intention. Instead, we want to highlight and develop the claim that in all cases of communication, wherever they fall on the meaning/showing continuum or the determinate/indeterminate continuum, the intended import is achieved in the same way: by making mutually manifest one's intention to make an array of propositions manifest or more manifest to the audience.

Consider the version of Robert's story where he says to Susan, "Let's go for a walk" and she responds by opening the window and showing him the poor weather outside. By responding in this way, she makes manifest to him her intention to make manifest or more manifest an array of propositions which are relevant to his proposal to go for a walk, and which have become perceptually salient as a result of her opening the window. Of course, these are not the only propositions that her behaviour has made more manifest: it has become more manifest, for instance, that the window can be easily opened, that the street is noisy, and that she doesn't want to go for a walk because of the bad weather. However, out of all these propositions that her behaviour has made more manifest, Robert is able to identify the array

of propositions that she manifestly intended to make more manifest. How? They are the propositions whose increase in manifestness makes her communicative behaviour relevant in a way she may have intended and expected.

In another version, Susan responds to Robert's proposal by saying, "The weather is really awful!" By replying in this way, she makes manifest to him her intention to make manifest the proposition that the weather is really awful. Her behaviour also makes manifest a variety of other propositions: for instance, that she has a sore throat, and that she doesn't want to go for a walk because of the bad weather. Again, Robert is able to identify the array of propositions she has intentionally made manifest: as before, they are the propositions whose increase in manifestness makes her communicative behaviour relevant in a way she may have intended and expected.

When Susan's response is verbal, the array of propositions she communicates can be partly characterised by enumerating some of its members. These include the explicature(s) of her utterance (in this case, that the weather is awful) plus any implicatures that she made it wholly manifest that she intended to communicate (in this case, that she doesn't want to go for a walk). However, the full array of propositions communicated by Susan cannot be enumerated by listing (or providing a procedure for listing) all its members. Of course, enumerating all the members of an array is not the only way to identify it. For instance, it can also be identified by description. The array of propositions communicated by Susan contains the explicature and the one clear implicature of her utterance, plus all those implications of her utterance that were to some degree manifest to her and that she expected and intended her utterance to make more manifest to Robert in a way that would make her utterance optimally relevant.

When Susan responds non-verbally, by opening the window, the array of propositions she communicates can be characterised purely by a description: she intends to make more manifest to him those propositions which have become more manifest both to her and to Robert – i.e. which have become part of their mutual cognitive environment – as a result of her opening the window, and which are relevant because they imply an answer to Robert's proposal; the gist of this answer being that she doesn't want to go for a walk because of the bad weather.

Enumeration and description are not the only two ways in which an addressee may identify the array of propositions that a communicator manifestly intended to make (more) manifest to him. For instance, as a result of the communicator's behaviour, the addressee may experience a certain change in his cognitive environment, and identify this change, or part of

it, as something the communicator intended to cause in him and to have him recognise as what she intended to communicate. In this case, what is needed to identify the array is neither enumeration nor description, but merely metacognitive acquaintance.

Note that in talking of metacognitive acquaintance, we are not bringing onto the scene a new and unheard of kind of psychological awareness of the effects of other minds on our own. On the contrary, our awareness of the psychological effects that others have on us is a quite unremarkable aspect of our interactions with one another. We know it when our understanding of what others have in mind pleases us, angers us, shames us, makes us feel proud, and – less emotionally – makes us see things in a new light, makes us like or dislike things, makes us rethink the past and anticipate the future differently. We are often aware of the fact that a change of mind (whether or not we could spell out its exact content) was brought about by what we understood of the minds of others. What people do when they communicate is precisely to overtly reveal something of their own mind in order to bring about such changes of mind in their audience.

A central claim of *Relevance* and later elaborations of the theory is that, because of the Communicative Principle of Relevance, the addressee can intuitively identify the array of propositions communicated by an act of ostensive communication by following a task-specific ‘modular’ inferential procedure (Sperber and Wilson 2002). The same procedure applies whether the act is one of ‘meaning that’ or of ‘showing that’, and whether the intended import is determinate or indeterminate.

There are two specific propositions that any communicator, using any form of ostensive communication, makes it mutually manifest that she intends to make manifest: these serve as premises for the interpretation of her communicative act, rather than being part of that interpretation. The first proposition describes the particulars of the communicative behaviour of the communicator: for instance, that Susan opened the window in response to Robert’s proposal, or that Susan said “The weather is really awful!” in response to Robert’s proposal. Any theory of comprehension assumes that some such information is represented in the comprehender’s mind. Relevance theory claims that for any act of ostensive communication, there is a second immediately identifiable proposition that the communicator makes it mutually manifest that she intends to make manifest: the presumption that this act of communication is optimally relevant to the addressee (in a precise sense of ‘optimal’ defined in *Relevance*: 266–71).

Interpreting an utterance involves using these two propositions as premises (together with contextual information) in order to identify the array **I**. It is this identification that

constitutes the interpretation of the communicative act. A central claim of relevance theory is that this array is identified by following a path of least effort, and stopping when the resulting interpretation is such that the communicator could have expected it to satisfy the presumption of relevance automatically conveyed by that communicative act. On this approach, the intended import of a communicative act is not inferred on the basis of general maxims or principles, but on the basis of a presumption of the relevance of that specific act, which is communicated by the act itself without being part of its interpretation.

The addressee can identify the array **I** of propositions that an act of ostensive communication makes manifest in a variety of ways. We will consider cases of verbal communication in the next section. Here are some examples of non-verbal communication.

Peter asks Mary: “Did you bring your cell phone?” She answers by showing him her cell phone. Here, the array of propositions she makes it mutually manifest that she intends to make manifest to Peter may be a singleton: the proposition that she did bring her cell phone.

If Peter had said, “I’m sure you forgot to bring your cell phone,” Mary’s act of showing him her cell phone would have made manifest an array containing, on the one hand, two distinct propositions – that she did bring her cell phone and that Peter was wrong to be sure she had forgotten to bring it – and on the other hand a vague sub-array of propositions whose increase in manifestness would amount to shaming Peter for his lack of trust in Mary.

In a case of non-verbal communication with indeterminate import, as when Mary sniffs appreciatively and ostensively at the seaside air, the whole array **I** is vaguely identified as that array of propositions which makes (or is expected to make) the communicator’s behaviour optimally relevant to Peter. Suppose, for instance, that Peter was already appreciating the seaside air on his own: then her act may achieve relevance just by making mutually manifest that what is becoming manifest to each of them is mutually manifest.

## **7. What, if anything, remains of speaker’s meaning?**

One of our aims in this paper was to show that building an adequate theory of communication involves going beyond Grice’s notion of speaker’s meaning. Another was to provide a conceptually unified explanation of how a wider variety of declarative acts than Grice was concerned with – including both cases of ‘showing that’ and ‘telling that’ – are understood. We will end by considering where the resulting account leaves the notion of speaker’s meaning we began with, and what light, if any, it sheds on the rather fuzzy intuitions that Grice’s definition was designed to capture. In this last section, we focus on linguistic cases.

One intuitive distinction that Grice originally wanted his definition of speaker's meaning to capture was between 'deliberately and openly letting someone know' (by displaying direct evidence for the intended conclusion) and 'telling' (where all the evidence would be indirect). As discussed in section 2, the existence of a continuum of mixed cases involving both direct and indirect evidence (in different proportions and combinations) presents problems for this approach. In cases of 'showing that', either the evidence for the intended conclusion is not only direct but conclusive, or else some of the evidence (or at least some strengthening of the evidence) has to come from the communicator's intentions, and this cannot but lead to over-attributions of 'meaning' as defined by Grice. On the other hand, under-attribution of 'meaning' should occur with cases of 'telling that' where no evidence from the communicator's intentions is needed in order to accept the message, either because the logical structure of the utterance makes it self-confirming (as in Grice's example of a logical argument), or because the utterance refers to an object or event that provides conclusive evidence for the truth of the message (as in the utterance, "To open a champagne bottle, you can do this" where "this" refers to a demonstration of how to open a champagne bottle).

For philosophers of language and linguists who only want to use Gricean notions to discuss linguistic cases, one rather convenient, though un-Gricean, way to go to is to forget about the third clause of his definition of utterer's meaning, forget about his extended sense of 'utterance', and forget about his concern with providing a principled distinction between cases of meaning<sub>NN</sub> and cases of mere showing. Linguistic pragmatics is based on the distinction between what is linguistically encoded and what is verbally communicated. Grice illuminated our understanding of what is verbally communicated, i.e. of 'speaker's meaning' in the ordinary sense; so let's forget his attempt at a more ambitious theoretical definition, treat the first two clauses of his definition as necessary conditions characteristic of speaker's meaning, and make it a tacit rule only to study cases of verbal communication (or gestures like nodding that stand for verbal behaviour) for which this two-clause definition of speaker's meaning seems to pick out the intended phenomena.

We have given two reasons not to go that way. The first is that Grice was right to characterise, in a novel and ground-breaking way, a type of communicative behaviour – what we have called ostensive-inferential communication – that encompasses, but is not restricted to, verbal communication. The first two clauses of his definition of speaker's meaning (or better, the informative and communicative intentions proposed in relevance theory and inspired by Grice) do identify a fundamental form of human interaction, in the context of which, inter alia, verbal communication can be better understood and studied. The second

reason is that one and the same conceptual twist—starting from the assumption that the aim in all cases of human ostensive-inferential communication is to make an array of propositions (more) manifest—makes it possible to handle both the meaning/showing continuum and the determinate/indeterminate continuum in a unified way. Focusing solely on the study of verbal communication does not resolve the problem that this second continuum raises for the Gricean approach.

In verbal communication, as in non-verbal communication, we find that addressees exploit the full range of methods for identifying the array of propositions that the communicator intends to make (more) manifest.

In a case of ‘telling that’ with no implicatures, where the information the speaker intends to communicate is a single clear, paraphrasable proposition, the array is a singleton. An example would be the railway official’s reply “12:48” to the question, “When is the next train to Oxford?” In relevance theory, we define a notion of strength of communication which applies to individual members of a communicated array. A proposition is strongly communicated to the extent that it is strongly mutually manifest that the communicator intends to make this specific proposition manifest to the addressee. Any strongly communicated proposition falls unproblematically under the description ‘speaker’s meaning’. Thus, the proposition that the next train to Oxford leaves at 12:48 is strongly communicated by the railway official and is also a clear case of speaker’s meaning. Whenever the array of propositions **I** is a singleton, its single member is strongly communicated and is a prototypical case of speaker’s meaning.

In a case of ‘telling that’ with clearly identifiable explicatures and implicatures, the array can be identified by enumerating its members. Consider, for instance, the following exchange between new acquaintances chatting at a dinner party in London:

*Rob:* Do you live in London?

*Jen:* I live in Chelsea

Here, the relevance of Jen’s utterance depends on two clearly identifiable propositions: the explicature that she lives in Chelsea, and the implicature that she lives in London. These two are strongly communicated and are good cases of speaker’s meaning. Given that the exchange takes place in London, Jen may take the precision of the phrase “in Chelsea” to be relevant to Rob and not to carry any further implicatures. However, in appropriate circumstances it might carry such implicatures, and her tone of voice, looks, and so on (although not indispensable) might help to achieve these further effects. Since Chelsea is a

distinctly posh neighbourhood, Rob's social status would be a highly relevant contextual factor in interpreting Jen's utterance. If he were of a similar social status, her utterance might make weakly manifest that she is willing to share with him more personal information than he has requested, that she is not unwilling to see him again, and so on. The gist is clear, but no single proposition is strongly communicated or could be confidently described as part of Jen's meaning. On the other hand, if Rob were, say, a poor academic, then Jen's utterance, depending on the tone of voice, might weakly implicate that she belongs to a different and superior milieu and is not eager to deepen their acquaintance, or, especially if her tone of voice is apologetic, that she is aware of living in privileged circumstances and wishes he would not hold it against her, and so on. In all these cases, if Rob is at all savvy, he will correctly understand that something more than an answer to his question has been subtly communicated. The array **I** in this case is identified by its two strongly communicated members plus an awareness of some further cognitive effects that Jen was overtly intending to achieve by answering in the way she did.

It is all too easy for pragmaticists simply to ignore these weak effects and implicatures, and concentrate on strong implicatures that fit straightforwardly with the notion of speaker's meaning. However, as we have been arguing and will illustrate again below, there is a continuum of cases which should signal that the research is not quite on the right track.

The problem raised by such effects becomes harder to ignore when it affects not only implicatures but also explicatures. As we have argued (Sperber and Wilson 1998, 2008; Wilson and Sperber 2002), the content of implicatures and explicatures is inferred through a process of mutual adjustment whose goal is to produce an overall interpretation consistent with the addressee's expectations of relevance, and where the implicatures are warranted by the explicatures. The linguistically encoded sense of an utterance serves as a piece of evidence of the intended meaning, and provides a point of departure for constructing a contextually appropriate meaning that may be narrower or broader than the encoded meaning, or overlap with it, or even be identical with it, this latter outcome being neither preferred nor arrived at 'by default'.

Recent work in lexical pragmatics confirms that most encoded concepts are adjusted, or modulated, in the course of the interpretation process (Carston 1997, 2002; Sperber and Wilson 1998, 2008; Wilson and Carston 2007). Here is an example where the explicature contains an ad hoc concept constructed for the purposes of that particular interpretation:

*Mark:* We can't afford La Cantina

*Pamela:* I've got money.

Pamela's utterance "I've got money" would be literally true if she had 50c in her pocket; but it would not be optimally relevant. In fact, if she only had 50c, it would have been more relevant to produce the literally false utterance, "I have no money". What Pamela's reply communicates to Mark is the relevant information that she is both able and willing to pay for a meal at La Cantina. This implicature would not be warranted by the literal interpretation of her utterance, hence the construction of an ad hoc concept, MONEY\*, whose presence in the explicature warrants the implicature. But what exactly does she communicate explicitly when she says "I've got money"? The explicature, although not very vague, is not that easy to spell out. Pamela is asserting more than simply that she has money: if it turned out that she has 50c, Mark could justifiably complain that she had not just misled him but lied to him. Nor does she *explicitly* communicate that she has enough money to go to La Cantina: this is a consequence, rather than a rendering, of what she explicitly conveys. Basically, what she is referring to is an amount of money such that she is willing and able to pay for a meal at La Cantina, an amount which cannot be less than what the bill is likely to come to, but which may be quite a bit more. Mark might know her well enough to figure out roughly what amount she means, or he might just defer to her as the 'expert' (without this being a case of deferential meaning!). But in any case, there is no word or expression in English, or in any meta-language used by semanticists, which denotes what Pamela has in mind, and which she succeeds more or less in communicating. In other words, Pamela's explicit meaning in this case has a certain vagueness both for her addressee and for the analyst, and while this does not compromise communication between her and Mark in any way, it does compromise a standard account of her speaker's meaning.

Elsewhere, we and others working in the framework have given many examples and analyses of ad hoc concepts: some, like "money" in the preceding example, involve narrowing the linguistically specified meaning, while others involving broadening it, and others overlap with it. Note that we have described the case of "money" as one of narrowing, but in most situations, it could be a case of both narrowing and broadening, and hence of overlap. For instance, Pamela could very well say "I have money" when she has only a few dollars in her wallet, but also a credit card. In all these cases, pinpointing a proposition that would constitute the speaker's meaning is difficult or impossible. On the other hand, all these cases can be easily described as making more manifest an array of propositions any of which would warrant the implicature. The speaker is encouraging the hearer to accept any



proposition from this array as quite probable, while not committing to one of them in particular. This is not quite what is understood by ‘speaker’s meaning’, then, but perhaps it might be close enough for some people to be tempted to idealise away the complexity and ignore our proposal about how to treat it.

These were cases of fairly strong communication. But as we have emphasised, there is a continuum between these and cases of very weak communication, where any conceivable paraphrase of the speaker’s meaning would be quite defective. This happens when both implicatures and explicatures are weak, as is typically the case with metaphors.

Suppose that in an idle chat among friends, someone tells you, “Freddy is a waste of space”. The idea that a person could be a waste of space has no clear literal sense, and you will have to construct an ad hoc concept WASTE OF SPACE\* in the course of the mutual adjustment process in order to arrive at an array of implicatures that satisfies your expectations of relevance (which in this case are themselves likely to be rather vague and unconstraining). What you will probably end up with is a general impression of Freddy, based on the explicature that Freddy is a WASTE OF SPACE\*, an ad hoc concept derived by adjusting the linguistically encoded meaning in the light of whatever information is available to you about Freddy, your friend and the relations between them. In this example, both the explicature and the implicatures are weak: neither is easily paraphrasable, and although your friend has succeeded in communicating with you, you may find it hard to say exactly what she meant. On the other hand, the description of what is communicated in terms of increasing the manifestness of an array of propositions can be developed without idealising anything away. It won’t inject into the description the kind of precision and crispness that some would feel more comfortable with, but the phenomenon itself lacks both precision and crispness.

Or, to conclude with the classic example “Juliet is the sun”, the explicature (one might say) is Juliet is the SUN\*, where SUN\* is an ad hoc concept whose meaning is (vaguely) specified by mutually adjusting explicatures and implicatures in order to satisfy expectations of relevance: the explicature that Juliet is the SUN\* must carry an array of implicatures which makes the utterance relevant as expected, and the sense of SUN\* must be such that the explicature does indeed contextually imply these implicatures. These implicatures are weak, and cannot be enumerated. Hence, the explicature that warrants these implicatures is itself weak. There is no paraphrase in an adequate metalanguage – or even in English used as such a metalanguage – that provides a plausible analysis or rendering of the speaker’s explicit meaning. Even adding starred concepts to the metalanguage (as someone might suggest) would not allow us to identify a proposition as the speaker’s explicit meaning, since what a

starred concept does in this context is to vaguely indicate a range of possible interpretations that are all made more manifest (i.e. more probable and salient) without any one of them being THE correct interpretation. Just as Romeo need not have intended any one of these propositions to be taken as his exact meaning, so the audience need not, indeed should not, aim to attribute any exact meaning to him.

So the intended import of “Juliet is the sun”, as of so many creative metaphors, is best described as an array that the audience identifies not by enumeration but by metacognitive acquaintance, by attributing to the communicator’s intention what they mentally experience. In general, what is needed for successful communication is that the addressee’s mind be changed in the way overtly intended by the communicator, i.e. that the addressee be now disposed or more disposed to draw the kind of inferences the communicator intended (or at least that the addressee should understand the communicator’s intention, even if he does not fulfil it). The communicator need not intend the addressee to make this or that specific inference; her intentions may concern only the general drift of the addressee’s inferences and remain quite vague, and so may the addressee’s understanding, without this amounting to a failure of comprehension. What is aimed at in such cases of weak communication is a degree of cognitive alignment, not a duplication of precise contents.

What, then, remains of speaker’s meaning? Cases in one corner of the bidimensional continuum we have described; cases that have held the attention of linguists and philosophers of language at a time when pragmatics was non-existent, underdeveloped or, more recently, ignored; cases that we have tried to show do not have the kind of unity and autonomy needed to constitute a proper object of theorising. Like the proverbial drunkard in the night looking for his glasses under the lamppost not because of any strong reason to believe that they were there, but because at least he could see there, students of language have stayed close to the lampposts of semantics and logic. The drunkard’s strategy need not be irrational. But after a while ... especially if there are glimmers of light around...

### **Acknowledgements**

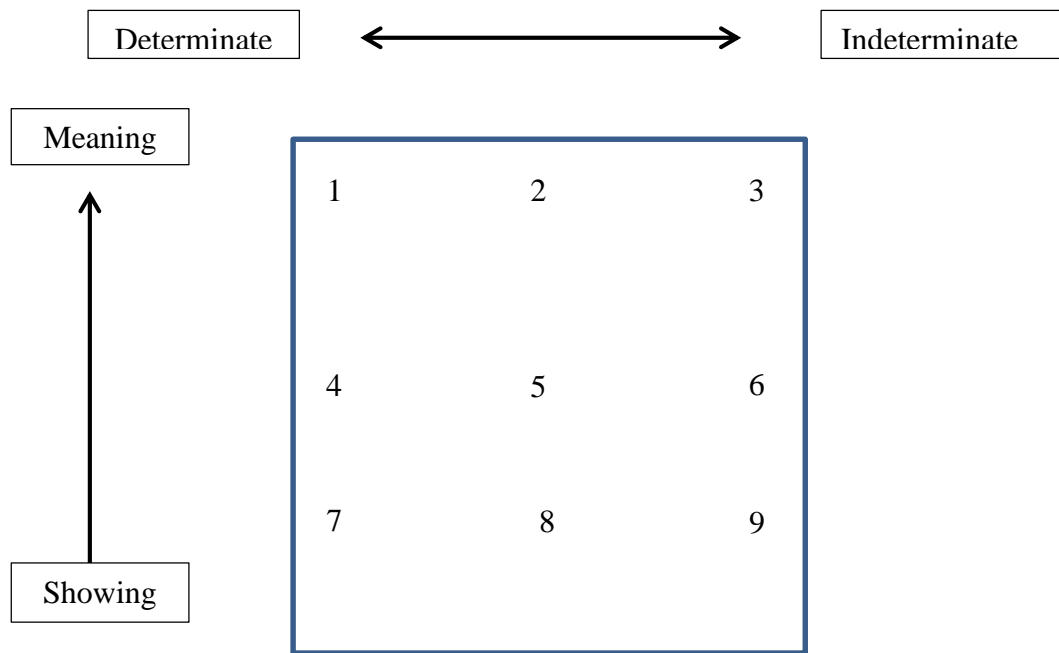
We are grateful to Dunja Jutronic and all the participants in the conference on Philosophy of Linguistics and Language, University of Dubrovnik, September 2014 for insightful comments and discussion during our presentation of this paper. We would also like to thank Nicholas Allott, Mark Jary, Stephen Neale and Tim Wharton for valuable comments on an earlier version.

## References

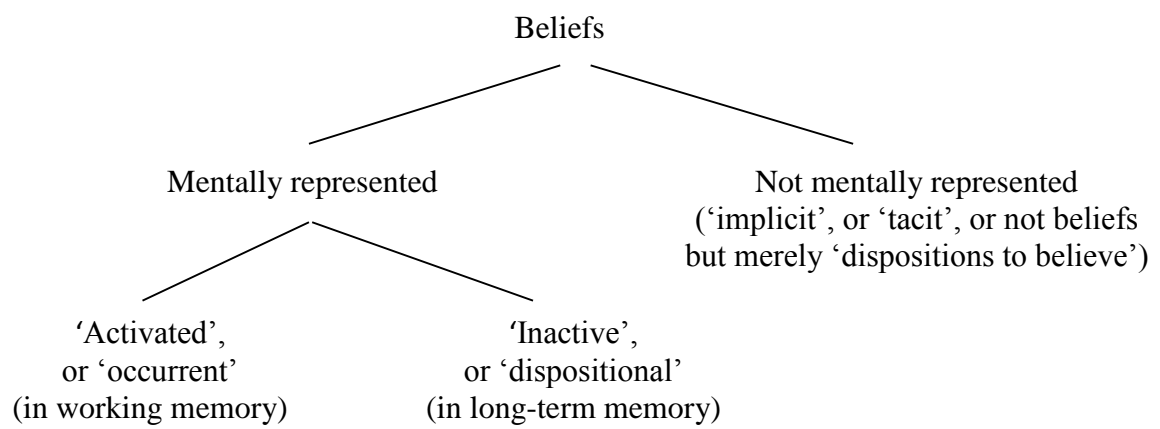
- Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509–59.
- Recanati, François. 1986. On defining communicative intentions. *Mind & Language* 1: 213–42.
- Schiffer, Stephen. 1972. *Meaning*. Oxford: Clarendon Press.
- Schwitzgebel, Eric. 2006. Belief. In *Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), ed. Edward N. Zalta. Available at: <http://plato.stanford.edu/archives/win2011/entries/belief/>
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. (2<sup>nd</sup> edition with a new Postface, 1995). Oxford: Blackwell.
- Sperber, Dan, and Deirdre Wilson. 1990. Spontaneous deduction and mutual knowledge. *Behavioral and Brain Sciences* 13: 179–84.
- Sperber, Dan, and Deirdre Wilson. 1998. The mapping between the mental and the public lexicon. In *Language and Thought: Interdisciplinary Themes*, eds. Peter Carruthers and Jill Boucher. Cambridge: Cambridge University Press: 184–200. (Reprinted in Wilson and Sperber 2012: 31–46.)
- Sperber, Dan, and Deirdre Wilson. 2002. Pragmatics, modularity and mindreading. *Mind & Language* 17: 3–23. (Reprinted in Wilson and Sperber 2012: 261–78.)
- Audi, Robert. 1994. Dispositional beliefs and dispositions to believe. *Noûs* 28: 419–34.
- Bach, Kent. 1987. On communicative intentions: A reply to Recanati. *Mind & Language* 2: 141–54.
- Carston, Robyn. 1997. Enrichment and loosening: Complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8: 103–27.
- Carston, Robyn. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Cavell, Stanley. 1965. Aesthetic problems of modern philosophy. In *Philosophy in America*, ed. Max Black. Ithaca, NY: Cornell University Press. (Reprinted in Stanley Cavell (1976) *Must we Mean what we Say?* Cambridge: Cambridge University Press: 73–96.)

- Clark, Herbert H. and Catherine Marshall. 1981. Definite reference and mutual knowledge. In *Elements of Discourse Understanding*, eds Aravind Joshi, Bonnie Webber and Ivan Sag. Cambridge: Cambridge University Press: 10–63.
- Dehaene, Stanislas. 1997. *The Number Sense*. New York: Oxford University Press.
- Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Neale, Stephen. 1992. “Paul Grice and the philosophy of language”. *Linguistics and Philosophy* 15: 509–59.
- Recanati, François. 1986. “On defining communicative intentions”. *Mind & Language* 1: 213–42.
- Schiffer, Stephen. 1972. *Meaning*. Oxford: Clarendon Press.
- Schwitzgebel, Eric. 2006. “Belief”. In *Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), ed. Edward N. Zalta.  
Available at: <http://plato.stanford.edu/archives/win2011/entries/belief/>
- Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. (2nd edition with a new Postface, 1995). Oxford: Blackwell.
- Sperber, Dan, and Deirdre Wilson. 1990. “Spontaneous deduction and mutual knowledge”. *Behavioral and Brain Sciences* 13: 179–84.
- Sperber, Dan, and Deirdre Wilson. 1998. “The mapping between the mental and the public lexicon”. In *Language and Thought: Interdisciplinary Themes*, eds Peter Carruthers and Jill Boucher. Cambridge: Cambridge University Press: 184–200. (Reprinted in Wilson and Sperber 2012: 31–46.)
- Sperber, Dan, and Deirdre Wilson. 2002. “Pragmatics, modularity and mindreading”. *Mind & Language* 17: 3–23. (Reprinted in Wilson and Sperber 2012: 261–78.)
- Sperber, Dan, and Deirdre Wilson. 2008. A deflationary account of metaphors. In *The Cambridge Handbook of Metaphor and Thought*, ed. Raymond W. Gibbs. Cambridge: Cambridge University Press: 84–105. (Reprinted in Wilson and Sperber 2012: 97–122.)
- Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25: 701–21.
- Vlach, Frank. 1981. Speaker’s meaning. *Linguistics and Philosophy* 4: 359–91.
- Wharton, Tim. 2009. *Pragmatics and Non-Verbal Communication*. Cambridge: Cambridge University Press.

- Wilson, Deirdre, and Robyn Carston. 2007. A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In *Pragmatics*, ed. Noel Burton-Roberts. Basingstoke: Palgrave: 230–59.
- Wilson, Deirdre, and Dan Sperber. 1988. Mood and the analysis of non-declarative sentences. In *Human Agency: Language, Duty and Value*, eds Jonathan Dancy, Julius Moravcsik, and C.C.W. Taylor. Palo Alto, CA: Stanford University Press: 77–101. (Reprinted in Wilson and Sperber 2012: 210–29.)
- Wilson, Deirdre, and Dan Sperber. 2002. Truthfulness and relevance. *Mind* 111: 583–632. (Reprinted in Wilson and Sperber 2012: 47–83.)
- Wilson, Deirdre, and Dan Sperber. 2012. *Meaning and Relevance*. Cambridge: Cambridge University Press.



**Figure 1**



**Figure 2**