RELEVANCE AND UNDERSTANDING

Deirdre Wilson
Universidad de Londres

1. Introduction

A certain politician - call her Margaret - is speaking to us on television. As she speaks, we are all noticing certain facts about her, interpreting these in the light of certain assumptions of our own, and coming to certain conclusions. In a sense, all of these contribute to our understanding of Margaret's speech; one might even say that they are all part of what she has communicated to us. In this paper, however, I want to think of communication, and understanding, in a rather narrower sense.

Notice first that not all the conclusions we draw were intended by Margaret. Take our conclusion that she is nervous: this is something she would have preferred us not to notice. Take our conclusion that she has no sympathy for the unemployed: this is something she would strongly dispute; if we drew it, she would feel misunderstood. Here, I will leave aside these cases of accidental information transmission and look instead at the domain of intentional communication and understanding.
Often, intentional communication involves a degree of manipulation or concealment. Like many politicians, Margaret, as she speaks to us, is doing her best to appear more intelligent, more sympathetic, more knowledgeable than she really is. These intentions can only succeed if they remain hidden: obviously, if we realise that Margaret wants us to think that she is nicer than she is, we are not going to be deceived. I want to leave aside these cases of covert communication and concentrate instead on a more basic, overt type of communication which we all engage in every time we speak.

In overt communication, there are no hidden intentions of the type described above. The speaker wants to convey a certain message, is actively helping the hearer to recognise it, and would acknowledge it if asked. During the question session after her talk, Margaret points to the back of the hall and says “I’ll take a question from the gentleman in blue.” In saying “the gentleman in blue”, she clearly intends to refer to someone, and clearly wants her intention recognised; to the extent that it is not, communication will fail. In this paper, I will take communication to mean overt, intentional communication, and understanding to mean recovering the overtly intended interpretation. I will try to show that understanding an utterance in this sense amounts to seeing its intended relevance: relevance and understanding are two sides of a single coin.

2. Understanding overt communication

Someone might claim that understanding an utterance is a simple matter of linguistic decoding. Margaret is speaking to us in English: it might be claimed that all we need to understand her is a knowledge of English. Virtually any utterance can be used to show that this hypothesis is wrong. There is a gap between knowing what a sentence of English means and understanding all that a speaker intends to communicate by uttering it on any given occasion. Communication and understanding involve more than mere linguistic encoding and decoding.
The examples that demonstrate the gap between sentence meaning and utterance interpretation fall into three main categories, corresponding to three main questions that the hearer of an utterance has to answer: (a) what did the speaker intend to say? (b) what did the speaker intend to imply, and (c) what was the speaker's intended attitude to what was said and implied? I will look at these questions in turn.

(a) *What did the speaker intend to say?*

Consider (1), which was taken from an advertisement for an employment agency that used to appear in the London Underground:

(1) If you're looking for a good job, we're offering a thousand a week.

Our knowledge of English alone will tell us that this advertisement has at least two possible interpretations: it may be offering a thousand pounds a week, or it may be offering a thousand good jobs a week. Our knowledge of English alone, however, will not tell us which interpretation was actually intended or understood. More generally, our knowledge of the language will tell us the range of linguistically *possible* interpretations of a vague, ambiguous or ambivalent utterance, but will not tell us which interpretation was actually intended on any given occasion.

In fact, this advertisement is quite interesting from a communicative point of view. It is what psycholinguists call a ‘garden path’ utterance: that is, an utterance on which hearers quite systematically get the wrong interpretation first, and have to correct it. Here, the first interpretation to occur to most English readers would be that they are being offered a thousand pounds a week, which is an awful lot of money - too much, in fact, to be handed out by advertising in the London Underground. Hence, this interpretation would have to be
rejected in favour of the less exciting interpretation on which what was being offered was merely a thousand good jobs a week.

Indeed, it is clear that these facts were deliberately exploited by the advertisers in order to attract the audience's attention. An advertisement which merely said "We're offering a thousand good jobs a week" would hardly have been worth a glance. An adequate theory of communication should explain not only the simple cases in which a vague, ambiguous or ambivalent utterance is correctly understood, but also more complex cases such as (1). Why is the first interpretation to come to mind generally the 'thousand pounds' one? On what grounds is it rejected? On what grounds is the 'thousand good jobs' interpretation preferred?

In the literature on communication, following the work of Paul Grice (1967/89), saying is generally contrasted with implying, or implicating. Every utterance is seen as communicating a variety of propositions, some explicitly, others implicitly. Saying is seen as falling on the explicit side. In order to discover what was said by an utterance (i.e. what proposition was explicitly expressed), the hearer must decode the sense of the sentence uttered, and then disambiguate any ambiguous expressions, assign reference to any referential expressions, restore any ellipsed material, and narrow down the interpretation of any over-vague expressions, all in the intended way. Thus, in order to know what Margaret intended to say in uttering the words "I'll take a question from the gentleman in blue", we would need to decide which particular man she had in mind. The resulting proposition determines the truth conditions of her utterance: it is the one that has to be true if her utterance is to be true. As these examples show, even the recovery of this explicitly communicated proposition is not a simple matter of linguistic decoding.

(b) What did the speaker intend to imply?

Sometimes, it is quite clear what the speaker intended to say, but less clear what she intended to imply. Consider (2), used by Mrs Thatcher in a BBC radio interview when she was still Prime Minister:
(2) I always treat other people's money as if it were my own.

Here, there is no problem deciding what Mrs Thatcher intended to say, but there is a problem deciding what she intended to imply. On the assumption that she treats her own money very carefully, (2) will imply that she treats other people's money very carefully; on the other hand, on the assumption that she spends her own money any way she likes, (2) will imply that she treats other people's money any way she likes, and so on. Different assumptions lead to different implications; the hearer's task is to identify the intended ones. Clearly, in this case the intended implication was that Mrs Thatcher treats other people's money very carefully, but how do we know this? More generally, how do we recognise the intended implications of any utterance?

In his *William James Lectures* (1967/89), Paul Grice introduced the term implicature to refer to the intended implications of an utterance. Some utterances have a few strong, highly salient implicatures; others have a broader, less determinate range. Thus, compare (3) and (4):

(3) a. *Peter:* Does Viv play cricket well?
   b. *Mary:* He plays for the West Indies.

(4) a. *Peter:* What will you do today?
   b. *Mary:* I don’t feel too well.

On the assumption that anyone who plays for the West Indies is a good cricketer, (3b) strongly implicates that Viv is a good cricketer, and the recovery of this implicature is essential to the understanding of (3b). (4b) has a broader and weaker range of implicatures, no one of which is essential to understanding. In uttering (4b), Mary clearly encourages Peter to think that she will be less energetic, less creative than normal, but she does not commit herself to any definite course of action. Here, it might be better to talk not so much of the intended interpretation as the intended line of interpretation, to account for the element of indeterminacy involved. In either case - whether the implicatures are strong or weak - they cannot be discovered by linguistic decoding alone.
(c) **What was the speaker’s intended attitude to what was said and implied?**

Sometimes, it is clear what the speaker intended to say or imply, but less clear what her attitude is to what she has said or implied. Consider a famous example from *Pride and Prejudice*. Elizabeth, the heroine, has finally agreed to marry Darcy, and her sister asks her when she first realised she was in love with him. Elizabeth replies:

(5) I think it was when I first set eyes on his magnificent estate at Pemberley.

The question raised by Elizabeth’s utterance is this: are we meant to think she believes what she said? In his review of *Pride and Prejudice*, Sir Walter Scott took the utterance literally, and condemned Elizabeth (and Jane Austen) for being mercenary. Many later readers have assumed that Elizabeth did not believe what she said: that she was indeed making fun of the idea that one might fall in love with someone for his magnificent estate. The issue, in other words, is whether Elizabeth’s utterance was intended as ironical or not.

A similar issue arises at the level of implicature. Consider (6):

(6) a. *Peter*: Is John a good cook?
   b. *Mary*: He’s English.

Given the reputation of English cooking, the most natural interpretation of Mary’s utterance in (6b) is that she intended Peter to supply the assumption that the English are bad cooks, and to conclude that John is a bad cook. But while she clearly intended to commit herself to the claim that John is English, it is less clear that she seriously intended to commit herself to the truth of the assumption that the English are bad cooks, and the conclusion that therefore John is a bad cook. Perhaps she was merely being playful, encouraging her audience to entertain the stereotype without actually endorsing it? Clearly, there is room for misunderstanding here.
In deciding on the speaker's intended attitude to the propositions expressed and implied, the audience has to answer the following sorts of question. Is she endorsing these propositions or dissociating herself from them; is she asserting that they are true, wondering whether they are true, perhaps wishing or hoping that someone will make them true? To a certain extent, these attitudes can be linguistically encoded (e.g. by declarative, interrogative or imperative syntax), but, as (5) and (6) show, in this aspect of interpretation as in any other, what is communicated generally goes well beyond what is linguistically encoded.

3. The nature and role of context

Understanding an utterance, then, involves answering three main questions: (a) what did the speaker intend to say; (b) what did the speaker intend to imply; and (c) what was the speaker's intended attitude to the propositions expressed and implied? It is obvious that context or background assumptions play a crucial role in answering these questions. By 'context' here, I mean not simply the preceding linguistic text, or the environment in which the utterance takes place, but the set of assumptions brought to bear in arriving at the intended interpretation. These may be drawn from the preceding text, or from observation of the speaker and what is going on in the immediate environment, but they may also be drawn from cultural or scientific knowledge, common-sense assumptions, and more generally, any item of shared or idiosyncratic information that the hearer has access to at the time.

Selection of an appropriate set of contextual assumptions is crucial to the understanding of (1)-(6) above. With (1), the audience must have access to the assumption that jobs paying a thousand pounds a week are not handed out by advertising in the London Underground. With (2)-(4), the choice of context is crucial again: once we know what contextual assumptions we were intended to use, the
intended implications follow by straightforward logical deduction. Finally, in (5)-(6), the difficulty of interpretation arises precisely because it is not clear what contextual assumptions we were intended to use: did Jane Austen in (5) mean us to assume that it is obviously ridiculous to imagine that one could fall in love with someone for his beautiful house, or did she mean us to assume that this was quite a reasonable thing to do?

Now if contextual assumptions affect the way an utterance is understood, then in order to recognise the intended interpretation, the hearer must select and use the intended set of contextual assumptions. Which adds a further question to our list of questions that the hearer has to answer: (d) what was the intended set of contextual assumptions? And in some ways, this is the most fundamental question of all.

In most writings on communication, while it is recognised that context makes a major contribution to understanding, the problem of how the intended context is identified is not seriously addressed. The assumption is that in normal circumstances only a single set of contextual assumptions could possibly have been intended. I want to argue that this assumption is inadequate, and that the problem of context selection is a genuine and serious one.

Imagine the following scenario. I am a keen club tennis player, and you know that I have recently begun playing with a new doubles partner. When we meet, you ask me what my new doubles partner is like, and I reply as in (7):

(7) He has much in common with John McEnroe.

At least for readers of the English tabloid press, the intended interpretation of this utterance will be immediately obvious. You are intended to use the contextual assumption that John McEnroe is extremely bad-tempered on court, and draw the conclusion that my new doubles partner is also bad-tempered on court. The question is why this is so.
Let's assume that our beliefs and assumptions about the world are organised in a sort of encyclopaedia in our minds under headings such as 'John McEnroe', 'tennis', 'doubles partner', and so on, and that in choosing a context for the interpretation of (7), the first place you will look will be under your mental heading for John McEnroe. Let's assume, too, that not all your beliefs and assumptions about John McEnroe are equally accessible, so that you don't pull out all of them at once. For most readers of the English tabloid press, as I have suggested, there is an immediate, very highly accessible assumption about John McEnroe that would come to mind in this situation: that he causes a lot of trouble on court. By using this assumption, and combining it with what was said in (7), you could derive the conclusion that my new doubles partner also causes a lot of trouble on court - which is the interpretation I have been assuming was correct.

Notice, though, that most people will have a lot more information than this stored under the heading 'John McEnroe'. You might know, for example, that John McEnroe is a very gifted tennis player, that he has a good serve-and-volley game, that he has played on the Centre Court at Wimbledon, that he is very rich, that he wears a headband when playing, that he is married to a film star, that he has a brother who plays tennis, that he enjoys rock music, and so on. By adding these assumptions to the context, you could derive a whole range of further implications: that my new doubles partner, like John McEnroe, is a very gifted tennis player, that he is very rich, that he has played on the Centre Court at Wimbledon, that he is married to a film star, and so on. What is there to stop you adding ever more contextual assumptions to the context, deriving ever more conclusions, and deciding that these were part of what I intended to imply? Notice, of course, that this is not what actual hearers would do.

This example suggests two important observations, which have to be taken account of in any adequate theory of utterance interpretation. First, it is clear that in interpreting (7) we do not assume that the speaker intended us to go on expanding the context indefinitely, deriving ever
more implications. We do look for some implications, of course; but what we appear to do is choose the minimal set of implications that would make the utterance worth listening to, and stop there. In the case of (7), we assume that the speaker's new doubles partner is bad-tempered on court in much the same way as John McEnroe is bad-tempered on court, and stop at that. An adequate theory of communication should explain why this is so.

Second, we do not - and could not - compare all possible interpretations of an utterance before deciding on the intended one. Intuitively, we do not do this; but, as this example shows, theoretically, we couldn't do it either, since for any utterance there is a huge range of possible contexts and possible implications, not all of which could conceivably be considered in the very short time it takes to understand an utterance. What we need, and what hearers seem to have, is some method of recognising the intended interpretation as soon as it presents itself, without necessarily considering any alternatives at all.

It should be clear by now that understanding an utterance involves considerably more than simply knowing the language. The class of possible interpretations is determined, on the one hand, by the meaning of the sentence uttered, and on the other by the set of available contextual assumptions. The hearer's task is to choose, from among this vast array of possible interpretations, the actual, intended one. In what follows, when I talk of the intended interpretation, I will mean the intended combination of explicit context, contextual assumptions and implications, and the speaker's intended attitude to these.

If the intended interpretation of an utterance is not recovered by decoding, how is it recovered? Paul Grice, in his William James Lectures, suggested an answer to this question. The intended interpretation is not decoded but inferred, by a non-demonstrative inference process - a process of hypothesis formation and evaluation - in which linguistic decoding and contextual assumptions determine the
class of possible hypotheses, and these are evaluated in the light of certain general principles of communication which speakers are expected to obey. According to Grice, speakers are expected to obey a Co-Operative Principle and maxims of truthfulness, informativeness, relevance and clarity; any hypothesis not satisfying these expectations can be automatically eliminated.

In our book *Relevance* (1986), Dan Sperber and I developed a theory of overt communication and understanding based on this fundamental idea of Grice's. In the next sections, I will explain the assumptions of relevance theory and apply them to a variety of examples. I will end by making some comparisons between relevance theory and Grice's earlier approach.  

4. Relevance theory

Relevance theory is based on a few very simple assumptions. First, that every utterance has a variety of possible interpretations, all compatible with the information that is linguistically encoded. Second, that not all these interpretations occur to the hearer simultaneously; some of them take more effort to think up. For instance, we saw with example (1) that the 'thousand pounds' interpretation is generally more accessible than the 'thousand jobs' interpretation, and with example (7) that, at least for most English tabloid readers, the assumption that John McEnroe is bad-tempered on court is easier to retrieve than the assumption that he has a brother who is a tennis player. As these examples also show, the order in which possible interpretations will occur to the hearer is at least to some extent predictable, though it is unlikely to be the same for all hearers at all times.

---

The third assumption is that hearers are equipped with a single, very general criterion for evaluating interpretations as they occur to them. In the case of (1), for instance, we know that the ‘thousand pounds’ interpretation will be rejected on some basis, and the ‘thousand jobs’ interpretation accepted. And the fourth, and final, assumption is that this criterion is powerful enough to exclude all but at most a single interpretation, so that having found an interpretation that satisfies it, the hearer need look no further: there will never be more than one.

The criterion proposed in Relevance is based on a fundamental assumption about human cognition. The assumption is that human cognition is relevance-oriented: we pay attention to information that seems relevant to us. Now every utterance starts out as a request for the hearer’s attention. As a result, it creates an expectation of relevance. It is around this expectation of relevance that the criterion for evaluating possible interpretations of an utterance is built. Different interpretations will be relevant in different ways: some will not be relevant at all; some will be fairly relevant; some will be very relevant. Which interpretation should the hearer choose? Clearly, the interpretation which best satisfies his expectation of relevance. To see how this criterion works, we need to know more about the nature of relevance and the expectation of relevance that every act of overt communication creates.

Relevance is defined in terms of contextual effect and processing effort. Contextual effects are achieved when newly-presented information interacts with a context of existing assumptions in one of three ways: by strengthening an existing assumption, by contradicting and eliminating an existing assumption, or by combining with an existing assumption to yield a contextual implication: that is, a logical implication derivable neither from the new information alone, nor from the context alone, but from the new information and the context combined. We claim that newly-presented information is relevant in a context when and only when it achieves contextual effects in that context, and the greater the contextual effects, the greater the relevance.
To illustrate these ideas, consider how the information in (8), given in tonight's weather forecast, might be relevant to you:

(8) It will rain in Paris tomorrow.

Suppose that you are going to Paris tomorrow, and already suspected that it was going to rain. Then (8) will achieve relevance by strengthening, or confirming, your existing assumption. Suppose instead that you are going to Paris tomorrow and were expecting it to be fine. Then, if you trust the weather forecast, (8) will achieve relevance by contradicting and eliminating your existing assumption. Finally, suppose that you are going to Paris tomorrow and have already decided to pack your raincoat if the forecast is for rain. Then (8) will achieve relevance by combining with this existing assumption to yield the contextual implication that you will pack your raincoat. All three types of contextual effect contribute to the relevance of (8), and the more contextual effects it achieves, the more relevant it will be.

Contextual effects, however, do not come free: they cost some mental effort to derive, and the greater the effort needed to derive them, the lower the relevance will be. To illustrate this idea, notice that (8) will seem more relevant to us if we really are planning to go to Paris tomorrow. In these circumstances, we will have no trouble thinking up an appropriate context, in which (8) will yield a satisfactory range of contextual effects, and hence be relevant to us. If we are not going to Paris tomorrow, we could no doubt still think up an appropriate context, but some effort of memory or imagination would be required. Intuitively, the greater the effort required, the less relevant (8) will seem to us.

The processing effort required to understand an utterance depends on two main factors. First, the effort of memory and imagination needed to construct a suitable context; second, the psychological complexity of the utterance itself. Greater complexity implies greater processing effort; gratuitous complexity detracts from relevance. Thus, compare (9a) with the longer and linguistically more complex (9b):
   b. It's raining in Paris and fish swim in the sea.

In circumstances where the hearer needs no reminding that fish swim in the sea, the extra linguistic complexity of (9b) will not be offset by any extra contextual effects, and will detract from the overall relevance of (9b) as compared with (9a).

The linguistic structure of an utterance is not the only source of psychological complexity. In fact, a linguistically simpler utterance may nonetheless be psychologically more complex. For instance, it is well known from psycholinguistic experiments that frequently-encountered words are easier to process than rarely-encountered ones. Thus, compare (10a) and (10b):

(10) a. I have no brothers or sisters.
   b. I have no siblings.

Although (10a) is linguistically more complex than (10b), most English speakers in most circumstances would regard it as stylistically preferable to (10b). The reason is that although (10b) is linguistically less complex, it contains the very rare word ‘sibling’, which generally requires more processing effort than the longer, but more familiar, ‘brothers or sisters’.

Relevance, then, depends on contextual effects and processing effort. The greater the contextual effects, the greater the relevance; but the greater the processing effort needed to obtain these effects, the \textit{lower} the relevance. The connection between relevance and understanding should now be clear. To see the intended relevance of an utterance, the hearer must identify the proposition and propositional attitude expressed, and combine these with the intended set of contextual assumptions to obtain the intended contextual effects; the intended set of contextual effects will include the intended contextual implications of the utterance, or what we have been calling its implicatures. To see the intended relevance of an utterance, then,
amounts to recovering the intended combination of content, context, attitude and implications. Relevance and understanding are two sides of a single coin.

The most basic assumption of relevance theory is that every aspect of communication and cognition is governed by the search for relevance. What is unique to overt communication is that, approaching an utterance addressed to us, we are entitled to have not just hopes but steady expectations of relevance. In the next section, I will argue that the expectation of relevance created by every utterance is precise enough, and powerful enough, to exclude all but at most a single interpretation, so that if we find an interpretation that satisfies our expectation of relevance, we can be sure that it will be the only one.

A word of caution here. Precisely because utterance interpretation is not a simple matter of decoding, but a fallible process of hypothesis formation and evaluation, there is no guarantee that the interpretation that satisfies the hearer's expectation of relevance will be the correct, i.e. the intended one. Because of mismatches in their memories and perceptual systems, the hearer may overlook a hypothesis that the speaker thought would be highly salient, or notice a hypothesis that the speaker had overlooked. Misunderstandings occur. The aim of a theory of communication is to identify the principles underlying the hearer's (fallible) choices. Relevance theory claims that the interpretation that satisfies his expectation of relevance is the only one that the hearer has any rational basis for choosing. To claim that a choice is rationally justified, however, is not the same as claiming that is invariably correct.

5. The criterion of consistency with the principle of relevance

The principle of relevance is the principle that every utterance (or other act of overt communication) creates an expectation of relevance. What exactly does this expectation amount to, in terms of effort and effect? Here, there is an obvious hypothesis: that what the hearer is
looking for is the *most* relevant interpretation: that is, the one that yields the greatest possible contextual effects in return for the smallest amount of processing effort. It is worth seeing why this hypothesis is wrong. In order to find the *most* relevant interpretation of an utterance, the hearer would have to consider and compare all *possible* interpretations; but as we saw when discussing the McEnroe example above, it is clear for both intuitive and theoretical reasons that hearers do not compare all possible interpretations of an utterance before deciding on the intended one.

In our book *Relevance*, Dan Sperber and I define a notion of *optimal* relevance which is meant to spell out what the hearer is looking for in terms of effort and effect:

**Optimal relevance**

An utterance, on a given interpretation, is optimally relevant if and only if:

(a) it achieves enough contextual effects to be worth the hearer's attention;
(b) it puts the hearer to no gratuitous processing effort in achieving those effects.

A word of explanation about each of these clauses.

On the effect side, what the hearer is entitled to look for is enough effects to make the utterance worth his while to process. In general, what this means is that he is entitled to expect more effects than he would have got from any other information that he could have been processing at the time. How much that is depends on what is going on elsewhere in his cognitive environment. Thus, suppose that someone walks into an important lecture and says (11):

(11) Ladies and gentlemen, I have to tell you that the building's on fire.
‘The building’ is a referential expression, and different assignments of reference lead to different levels of contextual effect. In the circumstances, the first hypothesis to come to the audience’s mind should be that ‘the building’ means the building where the lecture is taking place. Clearly, the utterance, on this interpretation, would achieve enough effects to be worth the audience’s attention: their minds would be immediately filled with thoughts of how to get out. Given that at a formal lecture the audience is supposed to be entirely absorbed in what the lecturer is saying, it is hard to see what other interpretation would achieve enough effects to justify the interruption; and in these circumstances, the interpretation just suggested is basically the only possible one.

It might be thought that in other circumstances the intended interpretation would be harder to pin down. Surely there might be several radically different combinations of content and context, each of which would yield enough contextual effects to make the utterance worth the audience’s attention? This is where clause (b) of the definition of optimal relevance comes in. Recall that we are talking about overt communication, where the speaker is anxious to avoid misunderstanding, and is actively helping the hearer to recognise the intended interpretation. Clearly, it is in such a speaker’s interest to make sure that there is no interpretation which is both easier for the hearer to construct than the intended one, and has enough effects to be worth his attention, since such an interpretation is likely to lead him astray. Clause (b) of the definition of optimal relevance, which excludes gratuitous calls on the hearer’s processing effort, covers this type of case: that is, it excludes the possibility that the hearer will be expected to recover, process and accept the wrong interpretation before lighting on the intended one. From clause (b), it follows that a speaker aiming at optimal relevance should try to formulate her utterance in such a way that the first acceptable interpretation to occur to the hearer is the one she intended to convey. From the hearer’s point of view, this clause has an immediate practical consequence. Having found an interpretation
which satisfies his expectation of relevance in a way the speaker might manifestly have foreseen, he need look no further. The first such interpretation is the only such interpretation, and is the one the hearer should choose.

I should note in passing that in order to be acceptable and comprehensible, an utterance does not actually have to be optimally relevant. Suppose that, as you are about to leave the house, I warn you that it is raining; as it happens, this is something you already know. In the circumstances, the proposition I have expressed will have no contextual effects and hence be irrelevant to you. Nonetheless, my utterance will be both comprehensible and acceptable as long as you can see how I might reasonably have expected it to be relevant. The actual criterion proposed in Relevance, then, is a criterion of consistency with the principle of relevance which takes account of this type of case:

Criterion of consistency with the principle of relevance

An utterance, on a given interpretation, is consistent with the principle of relevance if and only if the speaker might rationally have expected it to be optimally relevant to the hearer on that interpretation.

Vague as this criterion may sound, it makes one important prediction not matched by other theories. This follows from clause (b) of the definition of optimal relevance and its consequence that the first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance. Let us assume that, in interpreting an utterance, the hearer starts with a small initial context left over, say, from his processing of the previous utterance: he computes the contextual effects of the utterance in that initial context; if these are not enough to make the utterance worth his attention, he expands the context, obtaining further effects, and repeats the process until he has enough effects to make the utterance optimally relevant in a way the speaker could manifestly have foreseen. At that point, he has an interpretation consistent with the principle of relevance, and it follows that he should stop; or, at least, he is entitled to go on on
his own account, but is not entitled to assume that the speaker intended to communicate anything more. In other words, all the hearer is entitled to impute as part of the intended interpretation is the minimal (i.e. smallest, most accessible) context and contextual effects that would be enough to make the utterance worth his attention. Thus, the interpretation process has an inbuilt stopping place.

6. Some consequences of relevance theory

In this section, I will look at some practical applications of relevance theory to the analysis of a variety of examples. These will be grouped under two headings, to illustrate the two main strategies of analysis that relevance theory provides. Under the first heading fall analyses hinging on the assumption that the first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance; under the second heading fall analyses hinging on the assumption that any extra processing effort demanded will be offset by extra effects. Both strategies ultimately derive from clause (b) of the definition of optimal relevance, which excludes gratuitous demands on the hearer’s processing effort.

(a) The first acceptable interpretation is the only acceptable interpretation

This set of analyses all rest on the same point: that if an utterance has a highly salient (i.e. immediately accessible) interpretation which the speaker could have intended, then by clause (b) of the definition of optimal relevance this is the one she should have intended: she cannot rationally have intended to communicate anything else.

My first example was used by Jerry Katz in his book Semantic Theory (1972: 449-50). Suppose someone walks up and down outside the White House in America with a placard saying:
(12) George Bush is a crook.

He is then prosecuted for libelling the President of the United States. His defence lawyer argues that he was not intending to refer to the President, but was talking about a shopkeeper who had cheated him in his home town. As Katz says:

"It seems clear that such a defense would probably fail... The court would reason that the speaker must have known or can be assumed to have known that a national audience would inevitably take the occurrence of ['George Bush'] to refer to the President, and thus he ought to have employed a qualifying expression (e.g. 'who runs the grocery store in my neighbourhood') to make the statement that he says he intended to make." (p. 449)

In other words, if, in the circumstances, your utterance has a manifestly satisfactory and immediately accessible interpretation, that is the only interpretation you can rationally intend to communicate.

How does this follow from the definition of optimal relevance? Clause (b) of the definition says that the speaker should put the hearer to no gratuitous processing effort in achieving the intended effects. But it is clear that, in the circumstances described above, the speaker of (12) would have put his audience to some gratuitous effort if he intended to refer to anyone other than George Bush. He would have put them to the unjustifiable effort of first recovering, processing and accepting the wrong interpretation (i.e. the interpretation on which he was referring to the President of the United States), then wondering whether this was, in fact, the intended one, looking around for an alternative interpretation, and then, presumably, engaging in some further form of inference to decide between the two. Moreover, he could have spared his audience all this effort by reformulating his utterance in the way suggested by Katz, thus eliminating the unintended interpretation. It follows that the first interpretation consistent with the principle of relevance is the only
interpretation consistent with the principle of relevance: all other interpretations are disallowed.

My second example involves disambiguation. Suppose we are writing the plot of a television thriller, and I am giving you instructions about how the scenes should go. At one point, I say:

(13) In scene 10, when the police come in, the criminal makes a bolt for the door.

This utterance is ambiguous, with the two linguistically possible interpretations (14a) and (14b):

(14) a. When the police come in, the criminal runs for the door.
    b. When the police come in, the criminal gets out his tool kit and constructs a door-bolt.

In normal circumstances, the only legitimate interpretation would be on the lines of (14a). How can this be explained?

Relevancy theory suggests the following explanation. In the circumstances, the first interpretation to occur to the hearer should be (14a); moreover, this interpretation will be consistent with the principle of relevance: i.e. it will yield adequate effects for no unjustifiable effort in a way the speaker could manifestly have foreseen. Why should it be the first interpretation to occur to the hearer? Because it is based on a very stereotypical scenario. We have all frequently seen thrillers in which such a scene occurs; we should thus have easy access to a ready-made context in which to process the utterance on this interpretation: the police come in, the criminal runs to the door to escape the police, the police give chase, and so on. As a result, we should be able to achieve adequate effects for very little effort. By contrast, although we might all be able to think up a scenario in which the police arrive and the criminal, who happens to have a tool kit handy, sits down and hammers
out a door-bolt, it would take some effort of imagination to do so. Hence, an overall interpretation along the lines of (14a) will be easier to construct, and, once constructed, will prove satisfactory.

It follows from clause (b) of the definition of optimal relevance that a speaker, knowing this, is not free to intend an interpretation along the lines of (14b), and a hearer, knowing this, is not free to choose an interpretation along the lines of (14b). That is, on hearing my utterance in (13), you are not free to go out and shoot a scene based on (14b). In disambiguation, as in reference assignment, the first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance: all other interpretations are disallowed.

Exactly parallel arguments apply to the selection of intended context and implicatures. Thus, consider the following exchange:

(15) a. Peter: Would you like some coffee?
    b. Mary: Coffee would keep me awake.

In interpreting Mary’s utterance, Peter would normally be expected to supply the contextual assumption in (16) and derive the contextual implication in (17):

(16) Mary doesn’t want to be kept awake.
(17) Mary doesn’t want any coffee.

Notice that this is not the only possible interpretation. In certain circumstances - for instance if Peter and Mary are just about to attend a boring lecture - an interpretation along the lines of (18) and (19) might be both intended and understood:

(18) Mary wants to stay awake.
(19) Mary wants some coffee.

How does the hearer know which interpretation was intended? The answer again follows from clause (b) of the definition of optimal
relevance. If, in the circumstances, the contextual assumption in (16) is highly salient, and leads on to a satisfactory interpretation, then this is the only interpretation that the speaker is free to intend and the hearer to choose. Similarly, if, in the circumstances, the contextual assumption in (18) is highly salient, then this is the only interpretation that the speaker is free to intend and the hearer to choose. The first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance: all other interpretations are disallowed.

Returning to the McEnroe example, we can now explain why the hearer, having found a satisfactory interpretation based on the highly salient assumption that John McEnroe is bad-tempered on court, is not free to expand the context indefinitely, deriving ever more implicatures. By clause (b) of the definition of optimal relevance, as soon as this highly salient and satisfactory interpretation is discovered, all other interpretations are disallowed. The criterion of consistency with the principle of relevance thus explains the observation made earlier in connection with this example: that the minimal satisfactory interpretation is the one the hearer should choose.

(b) Extra effort implies extra effect

Let us look a little more closely at example (15) above, in which Mary says "Coffee would keep me awake", intending Peter to supply the assumption that she doesn’t want to stay awake and derive the conclusion that she doesn’t want any coffee. Would Mary’s utterance, on this interpretation, be consistent with the principle of relevance? If this were all she wanted to communicate, the answer would be ‘No’.

To see why this is so, we need to ask ourselves two questions, corresponding to the two clauses of the definition of optimal relevance: (a) could Mary have expected her utterance, on this interpretation, to achieve adequate effects? and (b) was there some other utterance (equally easy for Mary to produce) which would have achieved the
intended effects more economically? It seems clear that the answer to question (a) is 'Yes'. After all, by asking the question in (17a), Peter has indicated that a 'yes' or a 'no' answer would be adequately relevant to him. It seems equally clear, though, that if all Mary wanted to communicate was that she didn't want any coffee, she could have communicated it more economically by saying, simply, 'No'. Her utterance, on this interpretation, fails clause (b) of the definition of optimal relevance: it puts the hearer to some gratuitous effort.

It follows that if Mary was aiming at optimal relevance, she must have intended the indirect answer in (17b) to achieve some additional contextual effects, not achievable by the direct answer 'No'. Once alerted, we can see that this is so. By saying that coffee would keep her awake, Mary not only refuses the coffee but gives an explanation for her refusal - an explanation which would not have been communicated by the simple answer 'No'.

This example illustrates a very pervasive feature of utterance interpretation. Any element of indirectness in an utterance demands additional processing effort, and thus, by clause (b) of the definition of optimal relevance, encourages a search for additional effects, effects that a more direct formulation would not have achieved.

Returning to the McEnroe example, we can see that a similar sort of indirectness argument applies. I have suggested that the hearer of (7) is intended to supply the assumption in (20) and derive the conclusion in (21):

(7) He has much in common with John McEnroe.
(20) John McEnroe behaves badly on court.
(21) The new doubles partner behaves badly on court.

But if the only information the speaker wanted to communicate was that her new doubles partner behaved badly on court, why not say so directly? Why put her hearer to the additional effort of processing (7),
looking into his encyclopaedic entry for John McEnroe, retrieving assumption (20) and then performing a step of logical inference to arrive at (21)? It follows from clause (b) of the definition of optimal relevance that by putting her hearer to this additional effort, she must have intended to achieve some additional effect not achievable by saying simply (21). Here, the most likely line of interpretation is that she has in mind a particular degree and type of bad-temperedness: in other words, that she is encouraging her hearer to draw the conclusion in (22):

(22) The new doubles partner is bad-tempered on court in a similar way to John McEnroe.

And that, of course, is something that would not have been achieved by saying simply (21).

Metaphor provides a further type of indirectness argument. Consider (23) and (24):

(23) John is a lion.
(24) Bill is a donkey.

Many analysts of metaphor argue that these utterances communicate (25) and (26), respectively:

(25) John is brave.
(26) Bill is stupid.

Certainly, given stereotypical assumptions about lions and donkeys, these conclusions could be derived as contextual implications from (23) and (24). Within the framework of relevance theory, however, there is an indirectness argument to show that these analyses are inadequate as they stand. If all the speaker of (23) wanted to communicate was that John was brave, why not say so directly? Why put the hearer to the additional effort of processing (25), accessing the contextual assumption that lions are brave, and deriving (25) as a
contextual implication? A speaker aiming at optimal relevance must have intended to achieve some additional effects not achievable simply by saying (25): she might have intended to communicate, for example, not only that John is brave, but that he is brave in the way a lion is: the courage is physical rather than mental, depends on physical rather than moral strength, and so on. That is, she might have intended to communicate something more like (27):

(27) John is brave in the way a lion is brave.

And (23) might well have been a more economical way of achieving these effects than saying (27) directly. Parallel arguments apply to (24) and (26).

Indirectness involves making the hearer derive as a contextual implication something that could have been said directly. Not all arguments based on effort involve indirectness. Compare (28) and (29), for example:

(28) That was a stupid thing to do.
(29) That was a stupid, stupid thing to do.

It has often been noted that repetition can have an intensifying effect. Thus (29) might be understood as communicating something like (30):

(30) That was a very stupid thing to do.

Relevance theory suggests a natural explanation. By repeating 'stupid', the speaker puts the hearer to some additional effort. By clause (b) of the definition of optimal relevance, she must therefore have intended to achieve some additional effects, not achievable by use of the simpler (28). The most natural assumption - and hence the one favoured by considerations of optimal relevance - is that she thought the action described was stupider than would have been indicated by the
use of (28). On this interpretation, (29) would indeed have been equivalent to (30). FN2

7. Relevance theory and Gricean pragmatics

Grice's *William James Lectures*, delivered in 1967, offered the first systematic alternative to a code theory of communication and understanding. Focusing on the implicit aspects of communication, Grice argued that the implicatures of an utterance are not decoded but inferred, by a non-demonstrative inference process in which contextual assumptions and general principles of communication play an important role. His account of implicatures as beliefs that have to be attributed to the speaker in order to preserve the assumption that she has obeyed a Co-Operative Principle and maxims of truthfulness, informativeness, relevance and clarity, had instant appeal, provoked a flood of research, and is the starting point for most pragmatic theories today.

Grice’s insights left many questions unanswered. There were questions, in particular, about the nature and source of the Co-Operative Principle and maxims. Is co-operation essential to communication? Do speakers really aim at truthfulness, informativeness, relevance and clarity? What is relevance? Grice left this undefined. Where do the Co-Operative Principle and maxims come from? Are they universal? If so, are they innate? Are they culture-specific? If so, why do they vary, and how are they acquired? In our book *Relevance*, Dan Sperber and I set out to answer these questions. The resulting theory, sketched above, looks rather different from Grice’s.

There is a difference, first, over the role of the Co-Operative Principle and maxims. For Grice, the fundamental principle of communication is the Co-Operative Principle, according to which the

---

2. For further discussion of the effects of repetition, see Relevance chapter 4, section 6, pp 219-22.
speaker should try to make her contribution “such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange”. Grice assumed that every utterance, every conversation, has an accepted purpose or direction whose identification plays a crucial role in comprehension. This raises two questions: how is the accepted purpose of the utterance identified? and, once identified, how does it help with comprehension? Neither question has received a satisfactory answer.

Consider example (1) above (“If you’re looking for a good job, we’re offering a thousand a week”). The purpose of this advertisement is to attract people to the employment agency in question. As far as I can see, this purpose could be equally well achieved on either of the interpretations mentioned: how, then, does knowing the purpose of the advertisement help to choose between them? Consider the McEnroe example. The purpose of this utterance is to inform the hearer about the speaker’s new doubles partner. It could equally well achieve this purpose on either the minimal, correct interpretation, or any of the more expansive interpretations which we have seen are incorrect. How, then, does knowing the purpose of the utterance help with the identification of the intended context and implicatures? These questions have not been satisfactorily answered within the Gricean framework.

There is a more serious problem with the Co-Operative Principle. To the extent that the purpose of an utterance does play a role in comprehension, this merely adds a further question to the list of questions that the hearer has to answer: how is the accepted purpose of an utterance identified? Grice gives no answer to this. Like many theorists of communication, he seems to have assumed that the purpose of an utterance, like the set of intended contextual assumptions, is somehow given in advance of the comprehension process, or identifiable independently of it. In fact, it could not be identified by use of the Co-Operative Principle itself, on pain of circularity: to identify the purpose of an utterance by use of the Co-Operative Principle, one would already have to know it. Grice’s theory of
communication thus rests on the assumption that the purpose of an utterance is identifiable by some process that falls outside the scope of comprehension proper, and that is never satisfactorily explained.

Relevance theory suggests the following explanation. There is no Co-Operative Principle, and hence no circularity in assuming that the purpose of an utterance can be identified, where necessary, as part of the comprehension process. To the extent that the purpose of an utterance does contribute to comprehension, it is identifiable as a contextual assumption like any other, via the criterion of consistency with the principle of relevance. In this framework, there is room for a notion of purpose, but the real burden of explanation lies elsewhere.

Among the maxims, Grice sees truthfulness as the most important; relevance theory argues that there is no maxim of truthfulness, and indeed no maxims at all.\(^5\) Relevance theory is not a rule-based or maxim-based system. In this framework, relevance is fundamental to communication not because speakers obey a maxim of relevance, but because relevance is fundamental to cognition. As a result, the questions that arise in Grice's framework about the number of the maxims, their universality or culture-dependence, their acquisition, and the relative weight attached to each of them, do not arise.

A further difference between the two frameworks is over the role of maxim-violation. Grice listed a number of ways in which a speaker could violate the maxims: she could *opt out*, explicitly or implicitly, thus suspending a maxim; she could *covertly violate* a maxim, with intent to deceive; or she could *overtly violate* a maxim, thus creating an implicature. Although the mechanisms involved were unclear, the assumption that overt violation can create an implicature plays a crucial role in Grice's framework, and in particular in his account of metaphor and irony. This assumption has rarely been questioned (though see

Hugly and Sayward 1979 for excellent discussion). Relevance theory rejects it.

The principle of relevance is not a maxim: it is not a rule that speakers can obey or disobey: it is an exceptionless generalisation about what happens when someone is addressed. In such a framework, it makes no sense to claim that the principle of relevance can be overtly violated to create an implicature. How, then, do metaphor, irony, and the other phenomena that Grice analysed in terms of maxim violation, arise?

Sperber and Wilson argue that metaphor is simply a variety of loose talk. In a framework with no maxim of truthfulness, where speakers are not constrained to say only what is strictly speaking true, speaking loosely is often the best way of achieving optimal relevance. Hence, metaphor should arise naturally in such a framework. As we saw in section 6 above, the interpretation of metaphor involves an element of indirectness. This calls for extra processing effort, which, according to clause (b) of the definition of optimal relevance, must be offset by extra effects. In this framework, indirectness, with its resulting increase in processing effort demanded and contextual effects achieved, does much of the work that maxim-violation was supposed to do for Grice.⁴

A further difference between the Gricean approach and relevance theory is that whereas Grice was mainly concerned with the implicit side of communication, relevance theory has been equally concerned with the explicit side. Relevance theorists have looked in particular at the role of contextual factors in disambiguation, reference assignment and other processes that contribute, in Grice’s terms, to what was said rather than what was implicated: that is, to the truth-conditional content of utterances. Much work has been done on distinguishing explicit from

implicit communication, and truth-conditional from non-truth-conditional meaning; this seems to me to have been a particularly fruitful line of research.\textsuperscript{F5}

Having drawn attention to some of the differences between Gricean pragmatics and relevance theory, I would like to end by underlining what they have in common. Relevance theory rests squarely on Gricean foundations: Sperber & Wilson accept Grice’s view that the goal of pragmatic theory is to explain how the hearer recognises the overtly intended interpretation of an utterance; they acknowledge the importance of non-demonstrative inference in comprehension, and agree with Grice that general principles of communication play a major role in the inference process, though not, perhaps, in quite the way Grice thought.

\textsuperscript{5} See, for example, Blakemore 1987, 1992; Carston 1988; Wilson & Sperber forthcoming. Several papers on relevance theory are collected in Davis 1991; see also two recent issues of \textit{Lingua} devoted to recent work on relevance theory.
REFERENCES


Sperber, Dan & Wilson, Deirdre 1987 ‘Presumptions of relevance, in Behavioral and Brain Sciences. 10.4: 736-54.


Wilson, Deirdre & Sperber, Dan 1992 ‘On verbal irony’. Lingua.