

Comprehension Procedures and Non-Communicative Utterances

Hanno T. Beck
The Benjamin Banneker Center

1. Introduction

One of the primary contributions of relevance theory (Sperber and Wilson 1986/1995) is an account of how linguistic communication can succeed, even though the encoded portions of utterances¹ greatly underdetermine that which is communicated. Relevance theory aims to give an account with greater cognitive realism and more accurate predictions than a pure Gricean maxim-based approach.

According to relevance theory, under certain conditions (those of "ostensive-inferential communication"), a hearer can use a fast and easy comprehension procedure to arrive at a hypothesis of the speaker's meaning -- and a speaker can shape her utterances to take advantage of this phenomenon.

In this paper I will examine the conditions under which a hearer uses the relevance-theoretic comprehension procedure, and will explore some categories of utterance that lie at -- or beyond -- the limits of these conditions. I will consider various alternatives that might act as necessary and sufficient triggers for the operation of the relevance-theoretic comprehension procedure. Finally, I will look at some possible simplifications in formulating a heuristic-based answer to the problem of linguistic underdeterminacy.

2. Two or More Comprehension Procedures

Relevance theory claims that a hearer arrives at a hypothesis of the speaker's meaning by using a specific comprehension procedure. That procedure can be described briefly as "Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguation, reference resolutions, implicatures, etc.) in order of accessibility. Stop when your expectations of relevance are satisfied" (Wilson and Sperber 2004). It is by this process, the relevance-theoretic comprehension procedure (henceforth RTCP), that a hearer can receive an utterance that greatly underdetermines the "correct" interpretation, and, without sorting through a large (potentially infinite) number of possible

interpretations, can settle quickly on a single one. The reason that this one is normally the right one is due to the generalization called the 'presumption of optimal relevance' -- both speaker and hearer are tacitly aware that hearers follow a comprehension procedure along the lines described above, and therefore any intentionally communicative utterance can be reasonably expected to "mean" whatever that procedure yields.

This explanation is a regularity across contexts and languages -- all fluent hearers use the same comprehension procedure, under the right circumstances, according to relevance theory. I want to consider exactly what those circumstances are, and why those are the circumstances and not some others.

To address these questions, first we need to be clear about the separate roles of a comprehension procedure and that which launches the procedure. A comprehension procedure is logically distinct from the rules for its use. A theory can make claims about the nature of the comprehension procedure, and entirely different claims about the circumstances or criteria that put it into action.

If the RTCP is not used, a hearer still tries to make sense of the phenomena around him, whether those are linguistically encoded or not. There are two logical possibilities here. Either the hearer relies solely on general cognitive resources as a 'default comprehension procedure', or there is another special comprehension procedure available besides the RTCP. For the present, we will not argue for one possibility over the other, but will simply assume that there are no additional comprehension procedures besides the RTCP or the use of general cognitive resources. If our explorations find evidence for specialized work that must be done by a distinct comprehension procedure other than the RTCP, then general cognitive resources are not the only alternative to the RTCP and there must be a another way to interpret utterances. At the end of the paper we will offer a very speculative guess about this.

The Cognitive Principle of Relevance (Wilson and Sperber 2004) is that "human cognition tends to be geared to the maximization of relevance." When confronted with a phenomenon, if we pay attention to it at all, we do so while attempting to draw maximal meaning (positive cognitive effects) from it. If we are not obtaining sufficient positive cognitive effects to justify our processing efforts, then our attention moves to other phenomena in our environment. This is the default, or unmarked, comprehension procedure that we all use in response to a phenomenon that has our attention. The set of mental functions associated with this, I will call the Maximal Relevance Seeker, or MRS. This may for the present be taken as another name for "general cognitive resources." This comprehension procedure stands in contrast to the RTCP, which is built around a notion of optimal relevance rather than maximal relevance.

In most human activities, we use the MRS and we do not need to be conscious of it -- it is a natural part of our human endowment, it operates all the time that we are conscious, it is "always on."

The RTCP -- the comprehension procedure used to interpret phenomena that are communicative utterances -- is a special case, an exception to the MRS. The RTCP will apply only in a subset of the phenomenon-analyzing situations we find ourselves in -- namely, according to Sperber and Wilson, just those that are cases of ostensive-inferential communication, henceforth OIC.

The function of the RTCP, we might say, is to enable human communication to work -- some mechanism must enable us to infer, systematically and quickly, sufficient cognitive effects from a stimulus which explicitly encodes only a tiny, indistinct, impoverished hint. Later in this paper we will offer some speculations about the division of labor between the MRS and RTCP, with a view toward isolating what is unique to the RTCP.

3. The Activation Criterion: Selecting a Comprehension Procedure

Some phenomena, especially if they are linguistically encoded, will be recognized as OIC and so will be interpreted with the RTCP, while others (including some linguistically encoded phenomena and most of the vast range of non-utterance phenomena) will not, defaulting instead to being examined via the MRS. Part of an individual's linguistic and social competence is knowing when to use the RTCP. However, we are not claiming that either of these comprehension procedures is something that can be turned on or off consciously -- each is used at a tacit level, below that of consciousness.

Proponents and opponents of relevance theory have devoted a lot of attention to the RTCP or its competitors, but somewhat less attention to the conditions of its use. And yet, for every case where someone entertains an utterance, part of the interpretation process will involve a decision, made unconsciously and automatically, whether to use the RTCP. If the marked procedure, the RTCP, is not chosen, then the interpretation will be performed using the unmarked procedure, the MRS.

How is this decision made? The criterion that is used -- and we expect this to generalize across people and languages -- I will call the activation criterion, or AC. Using the AC, one tests whether to activate the RTCP. Thus the AC is the determining factor of how the remainder of the interpretation process will be enacted. We want to explore the details of this test. We will focus our exploration on phenomena that are linguistically encoded utterances.

When I say that a decision is made, or a choice is made, between using the RTCP or the MRS in utterance interpretation, I do not, at least at this stage, intend to presuppose that there is a single, observable moment when the hearer's mind ponders and determines which way to go, by means of some mentally-represented formula. It is possible that what I am calling the AC is not independently

mentally represented, but is just a regularity or generalization based on the interaction of other, more general principles.

What we are seeking initially is a formulation for the AC that is descriptively adequate, consistent with our knowledge and intuitions. Only empirical evidence will indicate whether this is also how people's minds actually perform the task. Nevertheless, we can confidently say that people receive utterances and they do end up, in virtue of some mental process, applying either the RTCP or the MRS. The AC that I will propose at the end of this paper is a cognitively plausible mental operation, not a mere byproduct of other factors, but empirical research to support that is minimal at this time.

Relevance theory offers an explicit definition of OIC and so commits itself to a particular set of conditions under which the RTCP will be activated, and this gives us a starting point for our exploration. Outside of relevance theory, the need for some sort of AC does not seem to be well recognized. Other pragmatists, Bach and Harnish (1979) for instance, do not appear to address this matter directly. For them, the domain of intended linguistic communication is all that is of interest. A 'communicative presumption' is in operation in all utterances, with the only exceptions being "recitation, elocution lesson, quotation" (p. 88). Bach and Harnish acknowledge the possibility that a hearer will decide whether the communicative presumption is "operative in a given context" (p. 7), but make no further investigation or claim as to whether this decision must be made consciously or unconsciously, what factors serve as inputs to this decision, etc. The notion is also omitted from their psychological account of comprehension (pp. 240-266). However, the mere fact that they admit the communicative presumption is not always operational, means that there must be a process, somewhere, for deciding when it is in force. That process, to do its work, can be located nowhere other than in the mind of the hearer, and must be operable at the time of utterance interpretation.

Later, Bach and Harnish do ask precisely the right question, but they direct it to relevance theory rather than their own theory: "Every act of ostensive communication is supposed to communicate a presumption of its own optimal relevance. But how does one identify a stimulus as an instance of ostensive communication in the first place, so that its optimal relevance can be presumed?" (Bach and Harnish 1987)

In developing an answer to that question, let us note that there are actually two logical possibilities of what the AC might be like if it is based on OIC. If, on the one hand, we take there to be some sort of unconscious mental representation of the concept of OIC, then the hearer can directly test for it. In this case, Sperber and Wilson could have answered Bach and Harnish's question

brusquely, by merely stating that the hearer identifies a stimulus as an instance of OIC, by evaluating whether it meets the definition of OIC. Let us call that our first candidate AC:

AC, Version 1. Is the utterance a case of OIC? If so, activate the RTCP.

(Since the MRS is the unmarked option, we do not need to add "otherwise, use the MRS.")

On the other hand, and more plausibly, the hearer might not have a mental representation of OIC, and will activate the RTCP only based on cues or signals. One possible example of such a signal is when the hearer finds himself unable to draw sufficient relevance from an utterance via the MRS.

My reading of Sperber and Wilson (1987, 1986/1995) is that they go with this second option, the idea that a hearer does not possess an internalized definition of OIC, but activates the RTCP when led to believe that there is relevance in an utterance (and finding no worthwhile relevance via the MRS). So we have a second candidate AC:

AC, Version 2. Was my attention ostensibly seized by someone? If so, has the person offered an utterance? If so, can I get sufficient relevance from this utterance to warrant my processing it merely via the MRS? If not, activate the RTCP.

In their reply to Bach and Harnish, Sperber and Wilson (1987) suggest that Bach and Harnish's question is answered in Sperber and Wilson (1986/1995) Chapter 3, Section 6. To summarize the point made there, an ostensive stimulus grabs the hearer's attention and then offers little or nothing that is relevant unless the hearer figures that it must be intentionally produced, thus promising at least a minimally sufficient degree of relevance (p. 154).²

Bach and Harnish, in their own pragmatic theory, cannot avoid a similar stage of analysis. No matter how they choose to define intended linguistic communication, there will always exist the real-world problem of deciding what meets their definition and what does not. A sound approach simply must recognize the inconvenient reality that the human mind does not know ahead of time whether a given phenomenon is an act of intended linguistic communication. People do not turn the RTCP on and off at random -- it is used under describable conditions.

One might conclude from the paucity of discussion of the AC that it may be a trivial matter. Perhaps the class of non-communicative utterances is so small as to be insignificant, or perhaps the distinction between OIC utterances and non-communicative utterances is clear, sharp and simple. In the next section we will survey some categories of utterances that will show that the AC is not at all trivial, the class of non-communicative utterances is enormous, and the distinction between OIC utterances and non-communicative utterances is often hazy and disorderly. As a byproduct, we will cast considerable doubt on both of our candidate ACs presented so far.

4. Communicative and Non-communicative Utterances

Sperber and Wilson (1986/1995) write that ostensive inferential communication takes place when "the communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions I" (p. 63). Carston (2002, p. 378) gives an almost identical definition.

Some of the generalizations and findings of modern pragmatics, in my view, pertain to utterance interpretation unconditionally. That is, they are generalizations about the MRS as well as the RTCP. To illustrate, we will see some examples of utterances that are perhaps not interpreted via the RTCP, but nevertheless have various features requiring pragmatic analysis. This may raise some new questions and options concerning how the interpretation workload may be divided between MRS and RTCP.

Are we safe in saying that linguistically encoded utterance-tokens will trigger the RTCP if and only if they are cases of OIC? This is an idealized view that has a lot of appeal but it does not, in my judgment, fit the data. Let us assume that a hearer is present, and that, for whatever reason, the utterance is linguistically encoded and has the attention of the hearer. Then if we are looking for utterance-tokens that fail to be OIC, we might expect them to fail for such reasons as these, or combinations of these:

- problems with the speaker
- problems with the informative intention
- problems with the communicative intention
- problems with mutual manifestness

I intend to present some examples of utterances that are not, or not necessarily, taken as OIC -- these would therefore fail the AC and would be interpreted via the MRS by default. But we must note a difficulty that has to do with empirical evidence -- how can it ever be shown that this is what actually happens?

Whether the RTCP or the MRS is used to interpret an utterance, we have, after a short time, an interpretation. In many cases both methods would yield the same result. Moreover, because the MRS is "always on," we must admit that as soon as an RTCP interpretation is made, or even before its completion, the MRS is at work seeking to extract additional relevance from the utterance. And, of course, all these acts of utterance interpretation occur below the level of consciousness. As a result, it is extremely difficult to distinguish in any decisive way, which comprehension procedure was used in

an interpretation. At present I am forced to rely on intuitions and to make distinctions and arguments based on those.

This is not a hopeless situation, however. Rieskamp and Hoffrage (1999, 2006) design experiments to show which of several competing heuristics was actually used in the making of a decision. They admit that we can only infer from observable factors such as time consumed and final output, and indirect evidence of factors apparently taken into account and their sequence, when we judge what processes were actually used. Nevertheless, well-designed experiments can add to our understanding, and we are in the same situation here. The high speed and sub-conscious operation of utterance interpretation processes makes them difficult but not impossible to investigate. Let's proceed.

(A) Gleanings and Overheard Utterances

Sperber and Wilson (1986/1995, p. 178) offer an example of a linguistically encoded utterance that is not OIC -- "when an actor doing voice exercises is accidentally overheard."

I will term such utterances "gleanings." With gleanings, there is no communicative intention, no informative intention, no attempt at mutual manifestness, and not even an intended audience; in a sense there is not even a speaker, but only an utterance source that could as well be an automaton or a recording. Still, the hearer might be able to derive various positive cognitive effects from that utterance. The utterance itself might exhibit ambiguity, unassigned referring expressions, generalized conversational implicatures (GCIs), ellipsed material, and nonliteral use of language.

Gleanings do not need to be characterized narrowly. Even if Sperber and Wilson's actor were communicating ostensibly with someone, and even if the overhearing was deliberate rather than accidental, the situation would be no different for our purposes. The hearer is confronted with an utterance which, regardless of its origin, has now arrived on the doorstep.

I call this larger category "overheard utterances." This is probably the most common class of non-communicative linguistically-encoded utterance -- the hearer is not a member of the intended audience, but his attention is captured anyway.

To interpret an overheard utterance, the hearer cannot rely on any presumptions of optimal relevance, or an informative intention, or a communicative intention -- because even if these were present in the utterance as intended for the hearer (and the existence of "gleanings" shows that that is not always the case), nothing at all was intended for the mere overhearer.³

This class of utterances is practically unbounded, since any utterance-token may be overheard in one way or another. Here are a few examples:

- (1) I saw a bat near the bank.
- (2) They said he could become president.
- (3) Some of my cousins went there last summer.
- (4) I don't think you did.
- (5) Reggie is a brick.

Overheard utterances may be relevant, and if the hearer endeavors to interpret them, he experiences no roadblock when called to perform such pragmatic activities as disambiguation (1), reference assignment (1-5), understanding generalized conversational implicatures (3, 4), infilling of ellipsed material (4), and recognition of nonliteral language use (5).

It will normally be clear when an utterance is overheard and thus non-communicative for that hearer, but there can be cases where judgment is needed. Mistakes can be made. For example, Lois is on the telephone when Ted enters the room. Lois says "You owe me a hundred dollars, and I want you to pay right now!" Ted might interpret the utterance as a case of OIC meant for him, or merely as an overheard utterance -- and whichever he chooses, he might be wrong. So might Frank, the person at the other end of Lois' telephone line. Lois' utterance was a case of OIC for one person (strictly speaking, it could have been intended for both Ted and Frank, but that is unlikely); Ted's task is not to determine whether the utterance is OIC in the abstract, but whether it is OIC for him, Ted.⁴ The kind of decision that Ted explicitly tries to make here, is offered implicitly to each of us every time we entertain an utterance -- whether easy or difficult, we must trigger the RTCP or not.

One way to characterize gleanings and overheard utterances is to describe them as reaching the hearer without a proper speaker, communicative intention, informative intention, or mutual manifestness. If these utterances capture and hold the hearer's attention, they do it by providing relevance to the MRS without involving the RTCP.

The category of overheard/gleaned utterances is vast -- we hear and pay attention to them every day -- but they are neither ostensive nor communicative in the relevance-theoretic sense.

(B) 'Deficient' Utterances

Sperber and Wilson (1986/1995) cite filibustering as another case of non-communicative utterance. In a filibuster, "All the usual features of verbal communication are present and even salient, but for one: there is no attempt at optimal relevance" (p. 159). One can imagine hearers, if not aware that a filibuster is taking place, gradually drawing this conclusion on their own after making many unrewarding attempts at interpreting the speaker's utterances as if they were OIC. At some point, the notion that each succeeding utterance is OIC grows so weak that they are rejected as such. The

utterances may still, however, have value for the hearer -- he may learn details of history or other subjects upon which the filibusterer is declaiming, and he will make use of pragmatic tools in his efforts to understand. Although the utterances are ostensive, this is not sufficient for communication.

Garcia Murga (1998), who also discusses Sperber and Wilson's example of filibuster, names several other instances of non-communicative utterances: "internal" speech in one's own head; soliloquy and emotional outbursts (without an audience); and speech to pets or objects. However, such utterances involve trivial, nonexistent, or non-comprehending audiences, and therefore they cannot be of use to us here as potential examples of utterances involving a competent hearer, facing interesting challenges in deciding whether to trigger the RTCP.

We require examples where hearers are not so unusual. I would like therefore to take Garcia Murga's list and consider the reverse of two of his examples. Instead of X speaking to no person, or to a pet, let us consider cases where X is spoken to by no person, or by a pet.

Jody enters a room and in her head she hears a voice uttering any of (6)-(10). Or, Martin enters a room and a parrot "speaks" to him, uttering any of (6')-(10'):

(6)(6') Flying planes can be dangerous.

(7)(7') Do you want a cup of water?

(8)(8') Most linguists like pizza.

(9)(9') Even Oscar would.

(10)(10') I want you to be a pig.

Assuming that a parrot does not have sufficiently rich internal representations to enable it to understand what it appears to be saying, nor to have communicative intentions, are all hearers prepared to ignore parrot-speech as mere empty squawkings? Are all hearers prepared to discount voices in their own heads as meaningless?

In fact it is hard not to take such utterances seriously. They appear to be linguistically encoded, and they capture the hearer's attention, seemingly ostensively. Yet they are apparently produced without a full speaker, without informative intention, and without communicative intention. At best, some parrots may have a notion that certain phrases tend to win food for them, but in other cases even this connection will not be present. But even if we are confident enough to characterize a parrot's utterance as not OIC, what of a chimpanzee's utterance? What of the utterances of a madman, a very young child, a severely handicapped adult, or a computer?

As for voices inside one's head, who can say with confidence what the source or sources might be? We can perhaps regard utterances emanating from an individual apparently possessed by a spirit,

in this same light. In those cultures where spirit possession is not unusual, hearers must judge utterances as OIC or not, in circumstances that would baffle many of us.

Utterances (6)-(10) show that the presence of ambiguity, referring expressions, generalized conversational implicatures, ellipsed material, and nonliteral language use do not prevent interpretation via the MRS.

Du Bois (1993) gives an example of utterances obtained through divination using a Zande (plural Azande) system known as the poison oracle or *benge*. The utterances obtained through this process are similar to voices from nowhere, since they are held to come from no person nor a spirit. The oracle gives answers by killing, or sparing, a fowl that has been fed a small amount of poisonous *benge*. In this way it is able to answer yes or no to any proposal set before it.

When one obtains a "yes" or "no" answer to one's quandary, there is no denying that new, relevant information has been brought into the situation. Among the Azande, in some situations it would be regarded as disrespectful, or would bring a loss of prestige, for one not to consult the oracle (Evans-Pritchard 1937, p. 261). The opinion of the oracle is highly respected but not absolute, much like that of a village elder, but *benge* is not personified; it does not represent any spirit or god (Evans-Pritchard (1937, pp. 320-22) stresses this point at length). This cannot be intended communication, since there is no intender. (To complicate matters further, there are other tribes, such as the Sissala, who do personify divinatory results, as the utterances of dead ancestors (Mendonsa 1982).)

Although "deficient" utterances appear not to be OIC, they form a messy category, not easy for a hearer to judge in real-world situations.

(C) Pass-Through Utterances

Sperber and Wilson (1986/1995, p. 177) note that "a linguistic stimulus triggers an automatic process of decoding" and argue that decoding takes place prior to the process of understanding (it is not necessary that decoding must be finished and complete before any inference can begin).

This fact, the automaticity of decoding, can be exploited by speakers as a way of conveying meaning, another kind of non-communicative utterance: one whose purpose is fulfilled merely in the decoding. Here is an example. Suppose that John is searching for the marmalade. Mary knows that the marmalade is on the upper shelf, and utters (11):

(11) Upper shelf.

Carston (2002, pp. 63, 152) analyzes utterances of this sort as genuinely subsentential, not mere ellipses of fuller, proposition-encoding sentences. However, at the pragmatic level she considers only

one possibility -- such utterances are incomplete and "there is some indeterminacy regarding the conceptual material to be supplied in constructing the proposition communicated" (p. 153).

Certainly Mary, in uttering (11), might indeed be communicating a proposition along the lines of "I assert to you that looking for marmalade on the upper shelf is more likely to bring you success than continuing to look on the lower shelf." But there is another possibility here. Just as (11) might be treated as syntactically complete, so might it be pragmatically complete. Mary might have no intention of communicating a proposition at all. Rather, by exploiting the fact that decoding is an automatic process, Mary may be seeking only to cause John to decode the noun phrase 'upper shelf.' Unless there is a failure in the transmission of the stimulus, the hearer is certain to decode it -- "We automatically recover its semantic representation," write Sperber and Wilson (1986/1995, p. 177). Mary knows that this decoding will cause the concept 'upper shelf' to have greater saliency, to be more manifest in John's cognitive environment, and so increase the likelihood that John will search the upper shelf for the marmalade sooner than he would otherwise. I will call such cases "pass-through utterances," where the speaker's goal is achieved by the hearer's mere decoding activity, rather than by recognizing intentions.

Thus the utterance described in (11) is, I would argue, ambiguous: Mary is probably communicating ostensively, but she might be making a pass-through utterance. The distinction is perhaps subtle, but of significance since Carston's "construct a proposition" analysis assumes that (11) is a communicative utterance, while pass-through utterances, since they include no communicative intention, are non-communicative.

To add support to the claim that (11) is ambiguous between an ordinary and a pass-through reading, we can give scenarios where the possibilities stand in greater contrast. Adding nearly anything to the end of the utterance causes it to carry further independent content and thus extinguishes the possibility that it is non-communicative. So we have no ambiguity in (12) or (13):

(12) Upper shelf, you dummy.

(13) Upper shelf, for crying out loud.

In (14) or (15) however, we have, to my ear at least, a preference for treating the utterance as a pass-through utterance:

(14) Upper shelf. (spoken softly, perhaps with high but falling intonation)

(15) Upper shelf, upper shelf, upper shelf. (softly, with even intonation)

In practice, Mary will probably be aware that her utterance in (11) is likely to be taken as OIC, but if she is making a pass-through utterance, this will be unimportant to her -- whether Mary utters it, or

Irene utters it, or the voice on the radio utters it, is unimportant if the goal is merely to have John decode the noun phrase.

A less natural but much sharper case for pass-through utterances can be made by considering ones that cannot be communicative at all, due to the absence of a communicator. Imagine John, once again searching for the marmalade, but with no one else present. A radio broadcast in the next room, or a remark in some separate conversation in the next room, causes "upper shelf" to be uttered just loud enough for John to hear it. (It is not even necessary that he hear it consciously.) John, by decoding the noun phrase, still winds up with greater saliency for "upper shelf" and his search may be influenced by this. In this scenario, there is no ambiguity: the utterance of "upper shelf" can only be a pass-through utterance, since there is no communicative behavior directed at John.

The key feature of pass-through utterances is that they are made with no communicative intention. The utterer, in typical pass-through circumstances, truly does not care whether she, or someone else, is perceived as the source of the utterance, and is not relying on authority or any other characteristic to influence the interpretation. Her intention is not even an informative one. Rather, her intention is only for her utterance to pass through the interpretive process, and there to affect the hearer's cognitive environment directly.

Because of pass-through utterances, hearers, when confronted with an example such as (11), have no automatic guarantee that it is a case of OIC.

(D) Proclamatives, Reminders

Pettit (1987) raises questions concerning the utterance in (16):

(16) William is coming.

Pettit acknowledges that on a standard pragmatic analysis, the utterer "manifests his intention to manifest the assumed fact that William is coming, and thereby hopes to manifest that fact itself." He goes on to suggest that (16) can be given an alternative analysis, however: "...the important possibility to register is that it conveys the assumption immediately and not by first presenting evidence of the speaker's intention to manifest the assumption." If the communicator does give evidence of a communicative intention, that is unimportant for it is merely "incidental to the success of his basic communicative enterprise."

Sperber and Wilson (1987) respond to this point, observing "As far as we can see, an utterance such as 'he is coming' provides no evidence for the conclusion that William is coming unless it first provides evidence that the speaker intends it to be so understood. In addition, the reliability of the speaker must be taken into account."

With examples such as (16), we indeed have an utterance that provides no evidence for the conclusion. Sperber and Wilson are not wrong, but there is a second possible scenario in play here. Evidence is not necessarily what the speaker is trying to provide in (16); rather, she may be trying merely to invite the hearer to entertain what (16) expresses. Providing evidence for (16) is one way, but not a necessary way, to do this.

When Sperber and Wilson suggest that speaker reliability is a factor, they are right so long as the speaker wants to be believed and is trying to achieve that, in part, by virtue of the utterance. But we can view Pettit's utterer as only seeking that Mary entertain the notion that William is coming. The utterer simply does not care whether he is recognized and/or evaluated as the source of the utterance. Utterance-tokens of this sort, where an informative intention is brought forward with an indifferent, or empty, communicative intention, I will call proclamatives. In proclamatives, the speaker is satisfied if the hearer interprets the utterance merely as if it were an overheard utterance.⁵

In many circumstances a proclamative will be misinterpreted by its hearer as including a communicative intention.⁶ These utterances are ostensive but lack a communicative intention. Proclamatives make no call for further utterances, and are expressed with no particular explicit communicative wish by the utterer. They are often spoken in a "take it or leave it" context. The utterer is not concerned to establish her authority, but is satisfied that her utterance is sufficient to achieve her (modest) goal.

In the class of proclamatives we have subsets of many types of utterance: proclamations, pronouncements, decrees, manifestos, declarations, many other performative utterances, recitations, reports, narrations, findings, advertisements, disclaimers, warranties, product guarantees, the outcomes of many forms of divination, and sometimes the final non-phatic utterance in a conversation. We might also include what we receive from checking the train schedule, or inquiring at the train station information booth. Du Bois (1993) cites consulting the dictionary, and referring to the rules of a game, as examples of this same sort. And unlike the utterance in (16), these latter examples -- train schedule, dictionary, game rules -- are not ambiguous between an OIC and a non-OIC reading. What we get from a dictionary, rulebook, or train schedule, is not a set of intentions.

Wilson (2003) mentions studies of autistic people, indicating they have impaired abilities to infer the intentions of others. I hypothesize that the autistic person's non-reliance on communicative intentions in utterance interpretation is similar to what the rest of us do when interpreting a proclamative utterance, where there really are no communicative intentions. An autistic person's world -- or that of a very young child -- is perhaps one where all utterances are analyzed as

proclamatives. Every utterance that arrives, is simply present. It can be accepted or rejected, but its reason for being there, why it was produced, is not or cannot be brought under useful scrutiny.

Millikan (1987) asks, "Perhaps there are some kinds of sentences that, at least when used in some kinds of ways, can be interpreted by Peter without his having to go through inference about Mary's intentions?" The answer is yes. Proclamatives would work in this way. In some cases it is impossible to infer intentions because there are none (at least none of the right sort); in other cases it is possible, but fruitless, to infer intentions because they contribute nothing to what is conveyed (much like pass-through utterances).

High authority is not absolutely necessary for an utterance to be proclamative -- even a low-authority speaker can utter a sentence S with the expectation that suspicious hearers will still entertain the possibility that S. In the real world, however, it is much more likely that proclamatives would occur in the case of a high authority speaker or an "authorized" source such as a dictionary or rulebook.

In many cases of divination, a ritual reinforces a supernatural authority and elicits an utterance from it -- a proclamative. For example, consulting the *I Ching* of China or the *hakata* dice of south-central Africa, involves acknowledging the authority status of the oracle, engaging in a formulaic act of random chance, and reading or reciting the outcome. The outcome is entertained -- whether it will be "obeyed" or in some other way integrated into the consultor's situation is an open question, but the new information is present.

Quotes, proverbs, and once again the outcomes of some divination processes, form another category of utterance where OIC is questionable. I will call these "reminders"; they form a distinguishable subset of proclamative utterances. In these cases, the utterer's intentions can come into play, but need not. When one says "A stitch in time saves nine" or "Practice makes perfect," one's own credibility or authority are not at stake. One is merely making a reminder, or citing common knowledge. Similarly, if a divination process results in one being told "You should not bring dishonor upon your clan," this is not news and authority is not really in play. These utterances are not bringing forward new material but merely raising the saliency and accessibility of material that already is manifest to the hearer. They increase the manifestness of a proposition/assumption, but not the strength with which it is held.

An activation criterion such as Version 2 above, will never trigger the RTCP if sufficient relevance can be drawn from the utterance without it; it would appear that no proclamatives, nor any intended non-proclamative that is capable of a proclamative reading, would ever trigger the RTCP under that AC.

(E) The Problem of Fiction

Sperber and Wilson (1986/1995) indicate that relevance theory is applicable to works of fiction, such as novels. Clark (1987) challenges this, and I want to revisit one of his remarks. Clark says that for Grice, "people talk ... as a means of doing things they can only do collectively." Grice (1975) proposes a Cooperative Principle for "talk exchanges," and offers his famous "conversational maxims" and discusses "conversational implicatures." Grice is careful to make his remarks pertain to conversation, rather than to communication in general.

I want to suggest that this is no accident.⁷ Conversation and talk exchanges have a here-and-now characteristic because the participants are jointly producing them. Fiction, on the other hand, breaks this connection between the participants by crystallizing one side in time, freezing its linguistic content into a book or other medium of reproduction. Reading a book is not the same as having a conversation, and the applicability of Grice's notions about conversation cannot be automatically assumed.

Sperber and Wilson (1987), however, do want the realm of pragmatics to include fiction. They write, concerning the works of Shakespeare and Melville, "These works are perfectly good cases of ostensive communication because they are, at the higher level, cases of showing rather than of 'saying that'" (p. 751). I object: something is a perfectly good case of ostensive communication if and only if it meets the definition of ostensive communication. The works of Shakespeare and Melville do not. Recall that one of the requirements -- not an option -- for OIC is that there obtain a state of "mutual manifestness" between the communicator and the hearer. Shakespeare and Melville, being dead, are unable to enter that relationship. Even if they were alive, still the mutuality of the relationship is impossible when one side is represented, not by a being with internal mental representations, but by a book. By definition, mutual manifestness describes a particular kind of shared cognitive environment, and people who are dead simply do not have cognitive environments.⁸ ("A cognitive environment of an individual is a set of facts that are manifest to him" and "A fact is manifest to an individual at a given time if and only if he is capable at that time of representing it mentally and accepting its representation as true or probably true" (Sperber and Wilson 1986/1995, p. 39).)

It may be suggested that when reading, say, *Moby Dick*, one imagines a living author with whom one has a relationship of mutual manifestness, and that in this way, fiction can still be regarded as OIC. However, such a proposal would overgenerate badly -- if we allow imaginary communicators, there is no limit to who could be imagined as the author of *Moby Dick*, and indeed any non-

communicative utterance at all could be converted into a case of OIC by fiat, simply by imagining an intending communicator.

Even with this amendment, one still has no basis for mutual manifestness; the best that can be achieved through this device is imaginary mutual manifestness.

Another possible response might be to say that fiction is OIC because the author, if he or she were present, would indeed have a shared cognitive environment with the reader. The author's story has merely been delayed in its transmission to the reader. However, this counterfactual strategy fails for the same reasons as the "imagining" strategy. It overgenerates, could be invoked in any case at all, and still does not fulfill the definition of mutual manifestness, which is not a "would be if only" state but a real-time state.

Written/recorded fiction does not meet the definition of OIC. Utterances separated from the utterer have a much less clear status than those in a real-time conversation.

Written/recorded fiction is not the only example of utterances that are separated in time from their receipt. Works of nonfiction, such as a scholarly article, also have this property. Should generalizations about conversation apply to these? Again, it appears that mutual manifestness does not obtain and cannot obtain when the communicator is separated in time from the hearer. (To be sure, there can be grey areas -- how much separation is too much? Even in fast real-time conversation there are tiny sub-second time delays.)⁹

What should we say of fiction where the author is not dead? I do not think that individuals, even unconsciously, take this into account. And in many cases, we do not know whether the author is alive, nor even who the author is/was. An account of understanding fiction that depends crucially on definitions affected by the living/dead state of the author, must be flawed.

Then can the same argument be extended to apply to nearly all written/recorded linguistic encodings, fictional or nonfictional, by authors alive or dead? Mutual manifestness does not seem possible in most of these cases. We may be tempted to conclude that what relevance theory is best at analyzing, and non-coincidentally what originally motivated Grice, the neo-Griceans, and the relevance theorists, is interpersonal real-time conversation. The analysis for conversation might not apply smoothly for other cases.¹⁰

(F) 'Continuous' Utterances

Some utterances fail to be clear examples of OIC because their "continuous" nature makes it difficult to assess not only who the utterer is/was, but what exactly the utterance event is. For example, consider the spine of a book, on which is written "Decline and Fall" and below that,

"Gibbon." If someone notices that book's spine, have they received an utterance? Did that utterance act take place at the present moment, or in 1966 when that particular book was printed and bound, or has it been uttered continuously that whole time? Is the utterance a case of OIC? To the last question it would seem the answer must be no, as there is no obvious communicator, nor an obvious communicative intention; there is therefore also no obvious mutual manifestness -- indeed, it would seem that these are fairly unobjectionable cases of something other than OIC. And yet, they are linguistically encoded expressions that require inference (even our rather plain example needs inference; "Decline and Fall" is not the full name of the book, nor is "Gibbon" the full name of the author).

We are not speaking merely of the spines of books. Road signs, computer icons and filenames, bumper stickers, labels on boxes and jars, political buttons, name tags, and countless others, are all examples of continuous utterances.

These utterances must be interpreted somehow and they can require the use of pragmatic tools such as disambiguation, reference resolution, etc.:

- (17) a. My son is an honor student at Westtown High School (bumper sticker)
- b. BRIDGE FREEZES BEFORE ROAD (road sign)
- c. GROOVED PAVEMENT (road sign)
- d. Inspected by 20 (paper inside a new pair of trousers)
- e. It's a Child, Not a Choice (bumper sticker)
- f. Green Party (bumper sticker)
- g. "P678" (bumper sticker referring to parking space P678 in the university parking lot)

Are these communicative, in the relevance-theoretic sense?

Imagine a road sign that reads "DETOUR, PROCEED LEFT." The viewer of that sign can easily understand its meaning without any reference to who, if anyone, intended it. This is a case of interpretation based on content, and context, and not on communicative intention. No reference to the installer of the sign, or the painter, or the sign designer, or the government agency responsible for maintaining the signs, or the legislators who passed the laws requiring the signs, or the lobbyists who demanded the legislation, or the fatcats who funded the lobbyists -- no reference to any of these possible candidates for "utterer" is necessary and none is calculated during the interpretation.

Continuous utterances are hints from which we infer what is inside, or what to do, or what is permitted/forbidden, or what has happened, etc. That, and the fact that we make these inferences swiftly and below consciousness, can make it tempting to categorize continuous utterances as cases of OIC. Sperber and Wilson (1987, p. 740) give an example of continuous utterances when they ask us

to imagine a botanical garden where the plants have labels. Addressing another topic -- whether communicative success requires uptake -- they claim "...only those labels which are invisible or illegible count as communication failures; all the labels which modify the cognitive environment of the visitors by making them *capable* of recognizing that the curators of the garden intended to inform them that this or that plant is called so-and-so are cases of communicative success."

I suggest that this situation can be analyzed differently. It is very doubtful that such utterances establish any sort of mutual manifestness. There seems to be no practical need for an intending communicator at all -- for regardless of that, the road signs and bumper stickers, etc., are interpreted and understood. One might even attempt to explore an analysis where these utterances are considered actually to be produced by the sign, spine, label, etc. as the utterer -- but that too would close the door on the possibility of mutual manifestness.

This completes our survey of some types of utterance where the hearer faces challenging work before he can activate the RTCP. What have we seen? The boundary between utterance-tokens that should be interpreted with one comprehension procedure versus another is often indistinct. Many communicative utterances yield sufficient relevance to the MRS and would never trigger the RTCP. Non-communicative utterances may embody characteristics that call forward such pragmatic tasks as the recognition of ambiguity; the recognition of needed reference resolution; the understanding of GCIs; the recovery of ellipsed material; and the recognition that part or all of the utterance may be interpreted nonliterally.

5. Utterance Interpretation Heuristics

Above, we named two candidates for the activation criterion. Both were built around a notion of OIC. In light of what we have seen in Section 4, we are now ready to look for alternatives.

Any AC must meet certain minimal conditions if it is to be cognitively plausible:

- a) it must operate below consciousness
- b) it must work quickly
- c) it should take as input an utterance-token
- d) it should yield as output a choice to use the RTCP, or not¹¹

In daily life, humans can and do perform this activity, without conscious effort and at high (sub-second) speeds. I want to suggest that the AC shows all the signs of a "fast and frugal" heuristic in operation (Gigerenzer *et al.* 1999).

I am not aware of relevance theorists openly claiming that a hearer literally reviews the definition of OIC every time he is confronted with an utterance (Version 1 of the AC). This would be

no more plausible than for a Gricean to insist that Grice's maxims are consciously consulted each time an utterance is interpreted. Rather, the claim that OIC is a criterion can be regarded in much the same way as a sensible approach to Grice's maxims -- these are generalizations describing what people do, but there is no claim that these particular formulations are represented as such inside people's minds.

Gigerenzer *et al.* (1999) suggest that a fast and frugal heuristic can operate on a very small number of considerations or "cues" and, if the outside environment corresponds structurally to what these cues capture, can yield excellent results.

We should look for a set of cues that can be resolved more easily than the definition of OIC. With regard to the AC, what would be some plausible cues? We might look for reliable signs that an utterance must be an intended communication:

AC, Version 3. Does the utterance embody one or more implicatures? Or irony? If so, activate the RTCP.

Implicatures, and irony, as defined by Sperber and Wilson, are intended. If a hearer can draw implicatures or irony from an utterance, then he should use the RTCP.¹²

However, these possible cues do not seem plausible with respect to timing. The comprehension process must already be underway for some time before it begins to grow clear that the likeliest interpretation of an utterance will embody PCIs or irony; yet common sense suggests that the AC plays a role early in the interpretive process, probably near the end of decoding and near the start of reference resolution and the like. It is not sound to rely on a cue that will not be available until well into the process, to determine in what way to begin that same process.

Another possibility is to consider whether an utterance includes "procedural" instructions along the lines analyzed by Blakemore (1992). If it does, that would be an indication that the utterance is an intended communication and should be analyzed with the RTCP.

AC, Version 4. Does the utterance contain "procedural" particles? If so, activate the RTCP.

In spite of the logic of that idea, in practice it is unlikely that the presence of Blakemorean procedural markers is sufficient to activate the RTCP. Procedurals might easily appear in overheard utterances, proclamatives, reminders, or continuous utterances, none of which seem to trigger the RTCP. Consider, for instance, utterances (18)-(20) if each was overheard, or received on the message slip in a fortune cookie:

(18) Sadly, not every cloud has a silver lining.

(19) Persistence pays off so don't give up, ever.

(20) Aim to live to 100 but don't give up all life's pleasures.

The typical hearer will encounter no difficulty with "sadly" or "so" or "but" in such utterances. And in any case, it is not a necessary condition that intended communicative utterances have at least one such marker, so we have to look further.

An utterance's linguistic properties are not likely to be good cues in a heuristic such as the AC. After all, any utterance might be communicative or might not, simply depending on whether the hearer is a member of the intended audience or merely overhearing it. So we will do better to abandon candidates such as Version 3 and Version 4, and look for cues in the circumstances and context of utterances, rather than in their content.

Much better, as possible cues for use in a fast and frugal decision tool, would be such criteria as:

AC, Version 5. Is there an explicit speaker directing this utterance to me? If so, activate the RTCP.

AC, Version 6. Is the communicator present here and now? If so, activate the RTCP.

AC, Version 7. Is there a situation requiring mind-reading (in some not-yet-settled sense of that term)? If so, activate the RTCP.

Here we are at least on the right track. However, candidates 5 and 6 will mistakenly activate the RTCP in cases of proclamatives. For those who believe that fiction triggers the RTCP, candidates 5 and 6 will also disappoint on that front. Candidate 6 will additionally fail to distinguish whether utterances are overheard. Opinions may vary but I think we can say that candidates 5 and 6 deliver mixed results at best.

Candidate 7 sounds like it has potential, but in its current form is badly underspecified. "A situation requiring mind-reading" must be explicated fully, otherwise it is just a statement of an intuition.

In an effort to capture the same basic insight with a simpler cue, we might try:

AC, Version 8. Are you in a turn-taking situation? If so, activate the RTCP.

This has an important element of truth, but even before we try to make precise the notion of turn-taking, I think there are ready counterexamples. For instance, in postal correspondence there is turn taking, but it can be doubted that the RTCP is always used.¹³

I wish to repeat my above-mentioned caution that empirically, it will be challenging to show convincingly when the RTCP is used and when the MRS is used. A person can use the RTCP to make

an interpretation and can then, almost at the same moment, use the MRS. Also, the waters are muddied by the fact that RTCP and MRS reach similar conclusions, and people are not conscious of which comprehension procedure they are using.

Another possible cue draws from the real-time characteristic of conversation:

AC, Version 9. Am I under time pressure? If so, activate the RTCP.

I will suggest that this is indeed the cue on which hearers rely. When we look over cases where the RTCP arguably does not operate, or where its operation is at least questionable, we find those situations that have less time constraint for the interpreter. Postal correspondence; reading works of fiction and indeed most any written work; hearing a proclamative utterance; pondering the output of a divination process -- none of these situations insist on immediate interpretation; rather, the hearer can take time and consider multiple interpretations, including ones that are quite unlikely to be intended by any speaker. The attitude toward possible interpretations is more one of collecting and less one of selecting. In the absence of time pressure, the hearer is free to pursue maximal relevance, and thus has no reason to use the RTCP.¹⁴

However, we have not by any means exhausted the supply of possible cues that might constitute the AC. Another good possibility as the primary cue used by hearers as a fast and frugal criterion for choosing whether to activate the RTCP in interpreting an utterance, is the need for a single final interpretation.

AC, Version 10. Must I come up with a single final interpretation? If so, activate the RTCP.

Situations where the hearer is obligated to come up with a single interpretation of an utterance tend to be the same as those where the hearer is operating under time constraints. And on the other side of the coin, overheard utterances, written/recorded works, proclamatives, etc., are characterizable not only as "low time pressure" for the hearer but also as "low pressure to settle on a single final interpretation."

These two alternatives, candidates 9 and 10, can perhaps be distinguished empirically, but it will not be easy. We would need to look for situations where time pressure is high while the need for a single interpretation is low, or vice versa.

As a possible case of the former, consider a driver who approaches a five-way intersection and observes a road sign on one of the alternatives, not identical to any she has seen before (perhaps she is driving in a foreign country). She was planning to follow that particular route away from the intersection. Under high time pressure, she must interpret the sign. However, there is no need to settle on a single meaning -- it is necessary only to gauge whether to continue down that road, or to take a different alternative. The sign itself might mean "proceed this way" but it could easily mean "road

damaged" or "stay out" or "no trespassing" or any number of cautions. If she interprets the sign as cautionary and proceeds down a different route, it matters not whether the precise phrasing on the sign indicated a damaged roadway, or the presence of man-eating plants, or any other detail.

In this situation, did the driver activate the RTCP? My intuitions are not clear. I hope at least that some examples along these lines might be further developed to help show how people are likely to react when under time constraints yet without the need for settling on a single final interpretation.

For the other possibility, cases where time pressure is low while the need for a single interpretation is high, postal correspondence is an example -- the hearer must settle on a single interpretation of the speaker's letter, in order to answer it, yet there is no special hurry to do so. In these cases, my intuition says that the MRS is in play, and therefore that the "time constraint" cue is a more accurate descriptor than the "need for a single interpretation" cue.

Considerable experimental effort would be needed to distinguish among these possibilities conclusively. And of course it remains possible that the actual cues are something entirely different, or some combination of those mentioned here, rather than a single cue.

We know from independent research that time constraints are a key dividing line; when under time pressure, subjects do have a tendency to switch to a fast and frugal heuristic (Rieskamp and Hoffrage 1999, 2006). The fact that people do change their decision strategies when under time constraints is an argument in favor of AC candidate 9.

While we do have Rieskamp and Hoffrage's empirical evidence for time constraints, we do not have any experimental evidence to cite for or against the "singleness of interpretation" cue. This is because the experiments conducted by Rieskamp and Hoffrage have used tests that involve subjects making judgments or choices under various conditions -- but there has been no variation where collecting possibilities without settling on just one has ever been an option. So we do not know whether there may also be experimental support for candidate 10; the experiments have not been designed nor performed.

The criterion that we are proposing as the AC -- "am I under time pressure" -- is not very specific. This will need to be addressed by further research, to assess the time parameters that people actually use, and to what degree the context may potentially stretch or compress these parameters.

Gleanings, overheard utterances, "deficient" utterances, proclamatives, reminders, written language, and continuous utterances generally carry little or no time pressure, and so with AC candidate 9 these would be predicted to be analyzed with the MRS, a prediction that I consider accurate. Pass-through utterances, since they contain no communicative intention, must on any account be analyzed without the RTCP and we would predict that too, because pass-through

utterances, which are completely interpreted as soon as they are decoded, would be unaffected by time pressure one way or the other. AC candidate 10 also would make these same predictions.

6. Workloads and Specializations

In this paper we have been talking about two comprehension procedures, the RTCP and the MRS, at a high level of abstraction. Recent works by Wilson (2003) and Sperber and Wilson (2002) are proposing that a dedicated cognitive mechanism performs specialized mind-reading tasks associated with pragmatics, as part of the RTCP. How radically should we expect the RTCP to differ from the MRS?

Gigerenzer *et al.* (1999) show that some fast and frugal processes are extremely different from "slow and burdensome" ones that seek the same goals. Theoretically there is no necessary structural similarity between the two. However, when presented with two processes that have similar inputs and outputs, Ockham's Razor suggests that we first assume that the two are much the same, and only hypothesize differences as dictated by the data.

Most linguists would accept that decoding is a process begun prior to the operation of the MRS or RTCP. In the absence of evidence, we would not want to say that there exist two separate decoding processes, one used with the RTCP and one with the MRS. This same argument, however, can be repeated for pragmatic activities such as ambiguity recognition, assignment of referents to referring expressions, etc., that we have seen can occur in utterances regardless of what mechanism is interpreting them. Unless we are compelled by other evidence, we should hypothesize that these activities take place unconditionally, and are not duplicated in the MRS and RTCP.

Whether he uses the RTCP or the MRS, a hearer performs such tasks as reference resolution, disambiguation, and the infilling of ellipsed material. A hearer draws on available information and cues from the immediate context, from encyclopedic knowledge, from memory, and makes deductions. What if we were to consider the radical hypothesis that these features common to the two comprehension modalities are, therefore, features of a single comprehension procedure? Because of its specialized nature, we will call it the UCP, or utterance comprehension procedure. We now have a more impoverished sense of the MRS, so let us rename it as merely the GCR, nothing more than general cognitive resources. And as for the RTCP, only a single task remains¹⁵

We agree with relevance theorists who propose that the RTCP is a fast and frugal heuristic; but we can move toward a revision of the RTCP as stated by Sperber and Wilson. If the interpretational work is actually performed via the UCP, then the only thing that we require from the RTCP is a search terminating factor. Again, in the spirit of proposing the minimum necessary to account for the

outward facts, I will suggest that the RTCP is just that -- a satisficing criterion built around the notion of relevance.

Satisficing is the establishment of a criterion and then the selection of the first item or option that meets or exceeds the criterion. Gigerenzer and Todd (1999), attributing the concept to Herbert Simon, define it like this: satisficing "takes the shortcut of setting an adjustable aspiration level and ending the search for alternatives as soon as one is encountered that exceeds the aspiration level" (p. 13).

The RTCP, when translated from descriptive terms into cognitive terms, is a satisficing process, namely a stopping rule. We can recast Sperber and Wilson's description of their RTCP into two parts. For the UCP, which operates unconditionally on utterance phenomena, "Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguation, reference resolutions, implicatures, etc.) in order of accessibility"; and for the RTCP, which is triggered under conditions of time pressure, "Stop at the first relevant enough interpretation". The actual semantic/pragmatic work is being done via the UCP; the RTCP need be nothing more than the stopping rule.

After all of this revising of boundaries and definitions, what we are left with is extremely close to relevance theory, with two important differences.

First, the UCP applies to a much broader range of utterances than merely those considered communicative in the relevance-theoretic sense. That is good, since we showed in Section 4 there are many utterances that seem non-communicative yet are interpreted by the hearer with standard pragmatic tools, and without hesitation or puzzlement.

Second, the notions of ostensive-inferential communication, mind-reading, and intentions no longer form the core of the theory. They still have roles to play, and they still occupy center stage when we restrict ourselves to interpersonal conversation, but those situations are a subset of utterance interpretation situations.

This second point would not be likely to win agreement from many relevance theorists. However, there are other developments pointing in the same direction. Breheny (2006), for an entirely different set of reasons -- to account for linguistic and cognitive behavior among the very young -- also argues that mind-reading is not so crucial a matter as has been believed. When there is a speaker and a communicative intention, those will be part of the context in which interpretation proceeds; in pragmatics, a lot of mind-reading activity does take place but it is not a necessity in every utterance.

To point out some possible areas for future research, I would like us to move toward constructing empirical tests that would help to choose between various formulations of the AC. As a starting point, we might look for utterances that are utterable under different levels of time constraint and see if we can thereby control the outcome of the AC.

Another area where we might be able to find some evidence where different possible ACs diverge, is mistakes. Mistakes are often useful in showing us things. Can we find cases of someone using the RTCP when they should not, or failing to use the RTCP when they should? We made a start with the example above concerning Lois and the telephone. A more systematic search for mistakes could prove enlightening.

7. Conclusion

In this paper, I have relied on many of the insights of relevance theory and have endeavored to challenge some others. Relevance theory, by offering an explicit definition of OIC, has made discussion of the AC possible, as we try to understand better the choices that a hearer must make during utterance interpretation.

Our survey of utterances that are non-communicative, or at least not sharp and clear cases of communication, has helped to highlight that Gricean and relevance-theoretic concepts of pragmatics apply best to cases of real-time interpersonal conversation (situations where time constraints are most clearly and consistently present). The further away from real-time conversation one gets, the less certain are these generalizations.

I have offered arguments showing the need for an activation criterion, suggested that structurally it is a fast and frugal heuristic, and have considered various candidates for it. The AC that I propose, namely the presence of time constraints, is supported by the limited evidence available at this time, but there are other good possibilities as well. The AC triggers a satisficing stopping rule and the first sufficiently relevant interpretation is taken.

The notion of a specialized mental module configured around mind-reading, for interpreting utterances of the intended communicative kind, is not supported by the findings and arguments in this paper; instead, I suggest that all utterances, OIC and non-communicative both, receive attention from a single specialized comprehension procedure, the UCP.

Finally, I described a very speculative realignment of relevance theory in order to capture a wider range of data. If something like it is true, then when a stimulus is received, the hearer begins to analyze it with the unmarked option, the GCR. If it seems to be an utterance, the UCP also begins working on it. And if there are time constraints, then the RTCP -- the satisficing stopping rule -- is also applied to the operation of the UCP. Mind-reading, ostension and communicative intentions still have important roles to play but no longer define the framework.

NOTES

¹ In this paper, I will use the terms *utterance* and *utterance-token* interchangeably, relying on the latter for emphasis. In both cases I have in mind any consciously-constructed expression; I do not assume the presence or absence of intentions. Moreover, when I refer to "utterances (6)-(10)" I hope it will be understood that I mean "utterance-tokens whose linguistically-encoded portions are transcribed in (6)-(10)."

² I will be arguing against both of these notions of the AC. Many communicative utterances can be quite relevant when interpreted via the MRS in just the same way as non-communicative utterances are, in which case they would never activate the RTCP. I believe that the RTCP is activated based on a different cue, one that I will try to show matches more closely our intuitions about which comprehension procedure is used with which utterance-tokens.

³ Let me be clear about the overhearer's task. Even though they cannot rely on any presumption of optimal relevance, informative intention, or communicative intention, this is not to say that an overhearer cannot use these factors in a *post hoc* interpretation. A hearer can imagine "what if this utterance had been communicated to me as the intended audience, how then would I interpret it?" In this situation, however, it is clear that the hearer is not using the RTCP. Rather, the hearer is imagining a nonexistent scenario, and surmising how he might have interpreted the utterance under those circumstances. This reconstructive action and the inferences based on it are performed via reasoning and the MRS.

⁴ Overheard utterances are at least likely to be cases of OIC for someone -- the intended audience. With gleanings, there is not even an intended audience.

⁵ Proclamatives are in this way similar to pass-through utterances, except that there is a speaker with an informative intention for the addressee; interpreting the utterance, rather than merely decoding it, is what the utterer desires, but he has no stake in steering that interpretation in any particular direction.

⁶ With typical intended communicative utterances the speaker deliberately conveys a communicative intention. With some utterances the speaker deliberately does not -- for example, Sperber and Wilson (1986/1995) describe Mary leaving the pieces of her broken hair-dryer lying around in order for Peter to notice and repair it, but hoping that he will not think she did this deliberately. The class of proclamatives lies between these two -- with these utterances the speaker is indifferent to conveying a communicative intention.

⁷ Pragmatists seem to have a joint agreement to ignore this point. Sperber and Wilson (1986/1995, p. 287, note 10) say "It is true that Grice's theory is put forward as a theory of 'conversation'. However, it has invariably been taken as a more general theory of verbal communication, and Grice has done nothing to correct that interpretation." It is high time that we faced up to this issue explicitly. A theory of conversation might not be adequate as a theory of communication, and converting it into one might not be easy nor straightforward. In this paper, the challenges posed to relevance theory concern precisely those areas where a theory of conversation looks uncomfortable or clumsy when extended to other, non-conversational forms of verbal communication.

⁸ In many cultures it would be claimed that spirits of the dead do indeed have cognitive environments; however, we need not consider that possibility here unless we wanted to propose an analysis of fiction that relied crucially on spirits of dead authors.

⁹ In my view, there is potential for looking at fiction in a different way. Perhaps instead of taking it to be OIC we might regard fiction, from the reader's point of view, as a collection of overheard utterances, packaged and distributed for the overall set of impressions it can give to the reader. The truth or falsity of those utterances (and hence most of the possible-worlds analyses attempted for fiction) are not pertinent. I would suggest that a work of fiction is close, structurally, to a (fabricated) recording of actual events, available for later playback. This would place it squarely in the category of overheard utterances, and so outside of OIC.

¹⁰ The view that intentions need not play any role in written/recorded works subsequent to their original creation, has some precedents. Du Bois (1993) cites favorably a 1974 article by Stallman, and a 1942 article by Wimsatt and Beardsley, where they argue that written works ought to be analyzed as self-standing entities, outside of any intentions that may have pertained to them. It seems likely, however, that those other authors were merely advocating ways to critique literature, rather than advancing claims about how the reader's mind works.

¹¹ Although we want to consider the AC as something like a universal, it is clearly plastic enough that it can vary across individuals and cultures. Wilson (2003) describes persons with Asperger's syndrome, and I believe they may be regarded as having an overly restrictive AC, interpreting most or all utterance-tokens via the MRS. On the other hand, in some so-

called "primitive" cultures, people are often described by outside researchers in ways that suggest an overly liberal AC; they are too ready to attribute intentions to spirits and inanimate objects, and thus to interpret events such as the sound of thunder or the creaking of boards as intended utterances..

¹² Since we have seen that GCIs can appear in utterances that are not communicative in the relevance-theoretic sense, we would need to restrict this criterion to just particularized conversational implicatures (PCIs) and cases of irony. Alternatively, we could redefine GCIs as something other than implicatures -- but not as explicatures, since those too are intended. If neither implicatures nor explicatures, I am not sure what other relevance-theoretic options are available for handling GCIs. Levinson (2000) suggests that GCIs occupy a separate level of analysis -- that of utterance-type as opposed to utterance-token. Utterance-type properties do not depend on intentions.

¹³ We can easily imagine a sentence that is many ways ambiguous, and a postal correspondent taking some time, even days or weeks, chewing over the multiple meanings.

¹⁴ Precisely what we mean by time constraints being "high" or "low" is something that empirical research will need to clarify. At least it is clear from Rieskamp and Hoffrage (2006) that there does appear to be a sea change, a natural boundary, where people choose different reasoning strategies based on time constraints.

¹⁵ One implication of this rearrangement is that the operation of the AC does not have to take place so early in the overall utterance interpretation process. If we rely on the UCP for reference resolution (or at least noting the likely candidates), disambiguation (or at least recognizing ambiguities, not necessarily eliminating them), and so on, it is possible that the AC comes into play later, or at many points. This is significant because it opens a door to a new possibility. I have argued above that the AC uses cues based on extralinguistic facts, such as whether the situation embodies time constraints, or whether a single interpretation is necessary. Before the interpretation process is underway, one is very limited in what one can say about the utterance itself -- but if the AC need not enter in until after some interpretation groundwork has already been performed, it could be possible that the AC does indeed refer to properties of the utterance as a cue. Although it would be unexpected, empirical research might find that the AC operates surprisingly late in the interpretation process.

REFERENCES

- Bach, K., and Harnish, R.M. (1979) *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.
- Bach, K., and Harnish, R.M. (1987) Relevant questions. *Behavioral & Brain Sciences* 10.4: 711-2.
- Blakemore, D. (1992) *Understanding Utterances*. Oxford: Blackwell.
- Breheny, R. (2006) "Communication and folk psychology." *Mind & Language* 21(1): 74-107.
- Carston, R. (2002) *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Clark, H. (1987) Relevance to what? *Behavioral & Brain Sciences* 10.4: 714-5.
- Du Bois, J.W. (1993) 'Meaning Without Intention: Lessons from Divination', in Hill, J.H., and Irvine, J.T., eds. *Responsibility and Evidence in Oral Discourse*, pp. 48-71. Cambridge: Cambridge University Press.
- Evans-Pritchard, E.E. (1937) *Witchcraft, Oracles and Magic Among the Azande*. Oxford: Oxford University Press.

- García Murga, F. (1998) "Communication, information, and relevance." *Revista Alicantina de Estudios Ingleses* 11: 75-83.
- Gigerenzer, G., and Todd, P. (1999) "Fast and Frugal Heuristics: The Adaptive Toolbox." In Gigerenzer *et al.*, pp. 3-34.
- Gigerenzer, G., Todd, P. and the ABC Research Group. (1999) *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.
- Grice, H.P. (1975) *Logic and Conversation*, in Cole, P., and Morgan, J., eds. *Syntax & Semantics 3: Speech Acts*. New York: Academic Press.
- Levinson, S. (2000) *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.
- Mendonsa, E.L. (1982) *The Politics of Divination*. Berkeley, CA: University of California Press.
- Millikan, R.G. (1987) What Peter thinks when he hears Mary speak. *Behavioral & Brain Sciences* 10.4: 725-6.
- Pettit, P. (1987) Inference and information. *Behavioral & Brain Sciences* 10.4: 727-9.
- Rieskamp, J., and Hoffrage, U. (1999) "When Do People Use Simple Heuristics, and How Can We Tell?" in Gigerenzer *et al.*, pp. 141-167.
- Rieskamp, J., and Hoffrage, U. (2006) "The Use of Simple Heuristics: Inferences and Preferences Under Time Pressure." Online document posted April 20, 2006; http://www-abc.mpib-berlin.mpg.de/users/rieskamp/Paper/Rieskamp_TimePressure.PDF.
- Sperber, D., and Wilson, D. (1986/1995) *Relevance: Communication and Cognition*. Second edition. Oxford: Blackwell.
- Sperber, D., and Wilson, D. (1987) Presumptions of relevance. *Behavioral & Brain Sciences* 10.4: 736-54.
- Sperber, D., and Wilson, D. (2002) "Pragmatics, Modularity and Mind-reading" *Mind & Language* 17: 3-23.
- Wilson, D. (2003) New directions for research on pragmatics and modularity. *UCL Working Papers in Linguistics* 15: 105-127.
- Wilson, D. and Sperber, D. (2004) Relevance theory. In L. Horn & G. Ward (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell.