C H A P T E R 27

RELEVANCE THEORY

FRANCISCO YUS

27.1 Cognition

Relevance theory (henceforth RT) is a cognitive pragmatics theory of human communication which was developed in the mid-1980s by Dan Sperber and Deirdre Wilson (henceforth S&W or, where appropriate, W&S) in their book, *Relevance: Communication and Cognition* (1986, 2nd edition 1995) but their earlier publications (e.g., W&S 1981; see section 27.3 below) also dealt with this theory, specifically comparing it to Grice’s cooperative principle. Since then, it has become a highly influential theory in today’s pragmatics and has been applied to many types of discourse and research areas within pragmatics (see Yus 1998a; 2006).

For S&W, human beings have developed an ability to maximize the relevance of the stimuli that they process. Since it is utterly impossible to pay attention to the entire barrage of information that reaches us, we have developed an inherent capacity to filter potentially irrelevant information and to focus our attention on what, in the current situation, is bound to provide cognitive reward. As W&S (2002a: 254) state, “as a result of constant selection pressure towards increasing

---

1 An extensive online bibliography on relevance theory, arranged in thematic sections and with links to more than 400 downloadable papers can be found in *Relevance Theory Online Bibliographic Service*, at www.ua.es/personal/francisco.yus/rt.html.

2 A stimulus is any input for mental processing. Stimuli can broadly be divided into verbal stimuli (e.g., an utterance) and nonverbal stimuli (e.g., a gesture).
efficiency, the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way”.

Filtering information which does not appear to be relevant (for example, when we do not recall most of the people who pass by us in the street but do remember those who, for some reason, stand out from the crowd) is indeed a typical mental activity oriented toward this relevance-seeking procedure, but it is also essential for human survival to identify underlying intentions and attitudes in the actions (communicative or otherwise) of those who are around us (for example when someone approaches us and we cannot help wondering what intention underlies his or her actions). Besides, human beings tend to select from context only the information that might be useful for obtaining interesting conclusions (contextual information is vast but we have developed a capacity for accessing just the right information that leads to interesting conclusions) and also combine new information with information already stored in their brain or which is accessible at that stage of interpretation (this is essential in human communication for obtaining interesting conclusions; see below). This inherent ability of humans to focus their attention on potentially relevant information is covered in the so-called cognitive principle of relevance.

27.1.0.0.1 Cognitive principle of relevance  Human cognition tends to be geared to the maximization of relevance.

This is a biologically rooted principle that is applied to all kinds of processing, including linguistic processing. This is reflected in the general objective of RT: to identify underlying mechanisms, rooted in human psychology, which explain how humans communicate with one another (S&W 1986/95: 32). Among the relevance-oriented tasks undertaken by the human mind, one of the most interesting ones is the human ability to combine contextual or accessible information with new incoming information to yield relevant conclusions, as in (1):

(1) New information (visual input):
  A yellow Mercedes is parked near our department.

(2) Information already available (from encyclopedic knowledge):
  a. Professor Smith, who supervises my thesis, owns a yellow Mercedes.
  b. Professor Smith usually takes the bus to university.
  c. Only when he intends to stay at university till late in the evening does he drive his car to university (since there are no late buses returning to where he lives).
(3) (Relevant) conclusion (inferred by combining (1) and (2)):
This evening I will be able to discuss with him at length how my thesis is progressing.

S&W claim that in a situation where (1) is processed, (3) would be relevant since it can only result from the combination of (1) and (2). A similar procedure also applies to linguistic communication (see section 27.2 below), specifically to intentional verbal communication. However, relevance is not only applied to external stimuli but also to internal mental representations, some of which are more prominent or likely to be entertained in the current context of interpretation. Consider the following example:

(4) The bell has just rung.
(5) a. Someone has rung the bell.
   b. The bell in my house has just rung.
   c. The person who is ringing is not a dwarf (he or she can reach the bell).
   d. There is no power failure in my building.
   e. The company providing electricity has not gone bankrupt.
   f. Nobody has stolen my ring.
   g. I have paid my latest electricity bill.

In situation (4), some thoughts are more accessible (more manifest in RT terminology; see section 27.5 below) and more likely to be entertained than others. In normal circumstances, (5a) and (5b) are the most likely thoughts. However the choice of thoughts is constrained by contextual information. For instance, in a context where there have been a lot of power failures recently, (5d) will then be more relevant and perhaps even more likely to be entertained than other thoughts which would be considered more accessible (manifest) in normal circumstances.

27.2 Communication

The biologically rooted capacity that human beings have developed in order to interact fruitfully with the surrounding world is also applied by S&W to communication, although, in this case, we are dealing with a highly sophisticated tool, language, which helps us transfer thoughts to one another. In a nutshell, whenever someone talks to us, we engage in a relevance-seeking inferential procedure which relies on the so-called communicative principle of relevance.

Commnicative principle of relevance:
Every act of overt communication conveys a presumption of its own optimal relevance.
When this principle is satisfied (normally, any time anybody addresses us, but also applicable to processing documents such as novels, news items, etc.), addressees undertake an interpretive task which aims at selecting the most appropriate interpretation from the range of interpretations that the utterance (or text) has in the current context. A stimulus has optimal relevance when two conditions are fulfilled. An ostensive stimulus is optimally relevant to an audience only if: (a) it is relevant enough to be worth the audience’s processing effort; and (b) it is the most relevant one compatible with a communicator’s abilities and preferences (W&S 2002a: 256).

On paper, hearers will follow the following general procedure:

(a) Follow a path of least effort in constructing an interpretation of the utterance (and in particular in resolving ambiguities and referential indeterminacies, in going beyond linguistic meaning, in supplying contextual assumptions, computing implicatures, etc.).

(b) Stop when your expectations of relevance are satisfied.

And for expectations to be satisfied, the selected interpretation should satisfy two conditions:

(a) An assumption is relevant to an individual to the extent that the positive cognitive effects achieved when it is optimally processed are large.

(b) An assumption is relevant to an individual to the extent that the effort required to achieve these positive cognitive effects is small.

On paper, new information is relevant when it reinforces the hearer’s assumptions about the world, when it contradicts and eliminates assumptions and, most importantly, when it combines with existing assumptions to generate conclusions (i.e., implications or implicatures) which cannot be obtained from either this new information or from the existing assumptions taken separately, but only from the combination of both. Consider the following example:

(6) Tom: So... Did you enjoy going to the cinema last night?  
Ann: John was also at the cinema.

For Tom to interpret Ann’s utterance correctly as an answer to his question (i.e., about whether she liked going to the cinema or not), he cannot simply interpret Ann’s words literally (which apparently have nothing to do with the question) but has to access contextual information (in this case encyclopedic information about Ann), for instance (7a–e) which, when combined with Ann’s words, will help Tom reach the intended interpretation (8):

(7) a. Ann has just gotten divorced.

b. Her ex-husband is called John.

Initially, S&W only mentioned “cognitive effects”, but in later publications a differentiation was made between those effects which are beneficial (positive cognitive effects) and those which are not. Needless to say, relevance is aimed at positive cognitive effects.
c. Now Ann and her ex cannot stand each other.

d. Whenever they come across each other they argue.

e. Ann gets depressed every time she argues with her ex.

(8) She didn’t enjoy going to the cinema last night.

Notice that this example does not differ too much from the one provided in (1)–(5) above. In (1) the new information was visual and it combined with stored encyclopedic information. In (6) the input is linguistic and it is also combined with encyclopedic information to get the right interpretation (i.e., the right conclusion). Crucially, (1) and (6) also differ in the role of intentionality. The former is an interpretation of unintentionally communicated information, whereas (6) involves obtaining an interpretation which Ann intentionally wants Tom to process. This is important because pragmatics does not undertake the study of information which reaches the person without a prior intention, although the cognitive mechanism to grasp relevant conclusions applies to both intentional and unintentional communication. Within this picture of intentional (specifically ostensive) communication speakers devise their utterances from among certain choices to code their thoughts, and hearers infer which interpretation, from among a choice of possible interpretations in the current context, is the one that the speaker intends to communicate. Wilson (1994: 44) summarizes these basic ideas of RT in four statements: (a) every utterance has a variety of possible interpretations, all compatible with the information that is linguistically encoded; (b) not all these interpretations occur to the hearer simultaneously; some of them take more effort to think up; (c) hearers are equipped with a single, general criterion for evaluating interpretations; and (d) this criterion is powerful enough to exclude all but one single interpretation, so that, having found an interpretation that fits the criterion, the hearer looks no further.

### 27.3 Grice and the Role of Intention

Most of the initial research by S&W on relevance was intended to acknowledge the importance of Grice in the history of pragmatics but also to criticize several points of Grice’s theory (see Grice 1975), specifically his emphasis on the need for a cooperative principle and its maxims to explain communication and also his dividing

---

4 For Grice, showing a cooperative attitude entailed the fulfillment of several maxims: maxim of quality (tell the truth), maxim of quantity (provide as much information as necessary), maxim of relation (be relevant), and maxim of manner (be brief, avoid obscurity of expression, don’t be ambiguous). When the speaker’s underlying cooperative attitude is taken for granted, the hearer can infer additional information (implicatures) from the fact that he or she is not following any of these maxims, supposedly for a reason.
line between explicitly and implicitly communicated information (see W&S 1981). A summary of S&W’s main criticisms of Grice’s theory is provided below:

(a) For Grice, understanding an utterance is a matter of constructing the best hypothesis about the speaker’s meaning. S&W agree with him on this. But Grice proposes the *cooperative principle* and its maxims as a means of evaluating alternative interpretive hypotheses. For S&W this *principle* is unable to do so.

(b) Similarly, the aim of a pragmatic theory is to explain how hearers identify the speaker’s intended interpretation. But Grice does not indicate how the *cooperative principle* would do that, so we cannot explain which interpretations are more likely to be selected by a hearer.

(c) Grice claims that the hearer should take into account the *cooperative principle* when selecting an interpretation but provides no insight into how this is done. Sometimes there are alternative interpretations to be chosen and various implications to be derived. A satisfactory pragmatic theory should be able to explain why hearers choose some interpretations and reject others. Grice’s framework does not do this.

(d) For Grice, pragmatic interpretation is an intelligent, inferential process which is based upon conscious reasoning. This is evident in the complex steps he suggested for the derivation of implicatures. But normally people are not really aware that they are inferring interpretations but, rather, engage in an unconscious and spontaneous mind-reading activity. S&W argue that an investigation of relevance helps us to understand why utterances raise the expectations they do, but it also leads them to reject Grice’s *cooperative principle* and its maxims and to construct an alternative theory within cognitive pragmatics.

However, S&W also acknowledge Grice’s important contribution to pragmatics, especially concerning the importance he gave to the role of intention in communication. Basically, for Grice, understanding an utterance involves recognizing the intentions underlying it. This is an evolved biological predisposition that humans also use in the interpretation of the nonverbal behavior of their interlocutors. What S&W do is to extend this view by proposing two types of intention. On the one hand, the speaker has an *informative intention*, the intention to communicate some information (“a set of assumptions” in the RT terminology). On the other hand, the speaker has a *communicative intention*, the intention to alert the interlocutor of his or her informative intention. Crucially, successful communication demands

---

5 “Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction” (Grice 1975: 45). The definition of his *cooperative principle* is as follows: “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.”
the fulfillment of both intentions (in this case it is called ostensive communication). This is particularly evident in the case of nonverbal stimuli. Consider the following situation:

(9) Tom and Ann are at a disco. At a certain stage during the night, Ann feels very tired and wants to go home. She sees Tom at the other end of the disco. He also sees her and waves at her. Since there is no way he can hear her at that distance, she moves her arms as if she was driving a car so that Tom can infer that she wants him to give her a ride home.

In situation (9) it is of the utmost importance for Tom to infer that Ann is intentionally producing her nonverbal behavior in order to inform him of something, otherwise Tom might think that Ann is simply engaged in some personal form of dancing. Only by inferring Ann's intentionality will Tom understand her correctly. In RT terms, only by inferring Ann's communicative intention can Tom understand the right informative intention ("I want you to take me home").

The valuable quality of verbal communication lies in how effectively it satisfies the speaker’s communicative intention. Indeed, when a person speaks to us we immediately infer that the person is willing to communicate some information to us, and therefore the communicative intention is immediately satisfied and we can focus our inferential activity on the message—informative intention—that the speaker intends to communicate. Verbal communication is indeed an invaluable tool for transferring thoughts.

### 27.4 Coding/Inferring

Traditional linguistic theories had a rather simplistic view of how people communicate information to one another. In short, for these theories all that speakers have to do is to code their thoughts into words and send them through a channel (e.g., the air, a book). Hearers are supposed to perform the same tasks but in reverse order: receive the words from the channel and decode them into the speaker’s thoughts. Duplication of thoughts between speakers and hearers was, from this point of view, the norm, rather than the exception. This is the code model of communication.

Within RT, by contrast, the picture is much more realistic and the emphasis is laid upon inference rather than on coding. Although thoughts have to be coded in order to be transmitted to other people, the information that speakers literally code is far more limited than the information that they really intend to communicate with this coded message. In other words, the utterance normally underdetermines
the information that is intended and this gap between what is coded and what is intended is filled by inference (see Carston 2002). 6 This is why the RT model of interpretation is called the inferential model. This model predicts that natural language sentences do not encode propositions but schemas for the construction of propositional forms. This idea is based on the evidence that we cannot possibly code literally the thoughts that we entertain, so instead of the uselessly lengthy (10b) a person should in normal circumstances utter a more relevant (10a):

(10) a. I think Susan put the book on one shelf in the sitting-room.
    b. I think Susan Thomas, my sister, put the book by Sperber that she bought two days ago on one of the shelves that are located in the sitting-room of this house, downstairs.

Under this underdeterminacy thesis (the claim that what people literally say is less informative than what they really want to communicate), there are two types of informational resemblance in human communication with gaps which have to be filled inferentially:

The interpretation that the speaker intends to communicate with his or her utterance

[is more informative than . . . ]

the literal meaning of what the speaker says, and this literal meaning

[is less informative than . . . ]

the interpretation chosen by the hearer.

Typical examples of “informational filling” are provided in italics in the following examples frequently found in the bibliography on this issue:

(11) I haven’t eaten.
    I haven’t eaten [this morning].

(12) It will take time to fix your car.
    It will take [longer than you’d expect] to fix your car.

(13) Everybody left early.
    Everybody [at the party] left early.

(14) There’s nothing on TV tonight.
    There’s nothing [worth watching] on TV tonight.

6 In this aspect, S&W also differ from Grice. For Grice, there is a very small gap between what is said and what is explicitly communicated, with reference assignment and disambiguation as the only pragmatic tasks undertaken to obtain the explicit interpretation. At the same time, he put all the inferential load into the derivation of implicatures, his famous coining. The work by RT analysts such as Carston (2002) has shown that obtaining explicit interpretations can be as demanding (or even more so), in terms of contextualization, as obtaining implicatures.
27.5 Mutuality

Within pragmatics it is commonly assumed that for communication to be successful, there has to be some information which is shared by both interlocutors, and on which they rely when communicating their thoughts. For example, for analysts of pragmatic presupposition, there is always some information supposedly shared by both interlocutors when an utterance is articulated and communicated successfully, as in (15):

(15) a. A: Where is Tom?
   B: There is a yellow BMW outside Sue’s house.

b. Pragmatic presupposition (supposedly shared): Tom owns a yellow BMW.

c. Implication (warranted by (b)): Tom must be at Sue’s.

Although it is intuitively certain that people do share information when they talk to each other, S&W reject this traditional notion of mutual knowledge because it generates an endless recursion (A knows that $p$, B knows that A knows that $p$, A knows that B knows that A knows that $p$, and so on) that prevents us from really assessing what is truly shared. Instead, they propose the notion of mutual manifestness (see S&W 1990). What is “manifest” is what one is capable of inferring or perceiving at a certain stage of interpretation, even if one has not yet done so. The sum of all the manifest assumptions is the person’s cognitive environment, which varies from one context to another.\footnote{Which means that manifestness is a matter of degree: in a certain context some assumptions will be more manifest than others, and hence more likely to be processed. In example (5) above, the thought that someone has rung the bell is more manifest than the thought that the person ringing is not a dwarf, even if both thoughts are manifest.} A set of assumptions manifest to several individuals constitutes their shared cognitive environment. This, again, changes according to different contextual parameters. When it is manifest to all the people sharing a cognitive environment that they share it, this is a mutual cognitive environment, made up of mutually manifest assumptions. Communication, then, is a matter of making certain assumptions mutually manifest to both speaker and hearer.

27.6 The semantics/pragmatics distinction

There is currently a lot of scholarly discussion on where to place the dividing line between semantics and pragmatics. Within RT the semantics/pragmatics
The context-free logical form is then enriched inferentially in order to obtain a fully contextualized interpretation (see section 27.8 below). In this picture, semantics would be in charge of the context-free semantic representation of the utterance, whereas the inferred interpretation would belong to pragmatics. In Blakemore’s (2002: 23) words,

The point of contact between semantics and pragmatics is at the interface between the linguistic parser, which receives input from linguistic competence and delivers linguistically determined semantic representations, on the one hand, and the inferential mechanisms which take these semantic representations as input for the computations which deliver the representations which the hearer takes to be representations of the speaker’s communicative intentions, on the other.

### 27.7 Relevance as a Cost–Benefit Procedure

Relevance is measured by hearers by following a cognitive cost–benefit procedure. All things being equal, the hearer will tend to select the interpretation, from the
range of possible interpretations of the same utterance in the current context,\(^8\) that satisfies the conditions of highest reward (positive cognitive effects) and least mental effort, although, as already suggested, hearers will normally be willing to devote extra cognitive effort if they are going to get additional effects.

As mentioned above, information which reinforces previous assumptions, contradicts and eliminates previous assumptions, or combines with previous or accessible assumptions to obtain conclusions provides the highest number of cognitive effects. Given a range of choices, the hearer is entitled to select the one providing the highest interest. But interest is constrained by effort. Wilson (pc) provides a clear example of how effort constrains the choice of an interpretation:

Imagine exactly the same information being presented, first in a clearly printed form; second as a faint photocopy; third as an illegible handwritten scrawl; fourth translated into a language you read only with difficulty. Each of these versions may have exactly the same cognitive effects for you, but each will require different amounts of processing effort. Although they carry exactly the same information, you will have to work harder to retrieve it from one input than from another, and this may affect your intuitions of relevance, and indeed, your willingness to attend to a particular input at all. More generally, an input may be more or less perceptually salient, more or less legible, more or less linguistically or logically complex, and may therefore cause more or less effort.

Notice, though, that this cost–benefit procedure does not imply that hearers invariably choose the most effort-relieving utterance. In practice, we will gladly devote cognitive resources to more effort-demanding interpretations if the eventual reward, in terms of positive cognitive effects, is worth the effort. The same applies to speakers when devising their utterances. An example of how speakers do not always opt for the most economical utterance is provided below:

(16) Ann: (1) Does Susan eat meat?
   Tom: a. She is a vegan.
   b. No. She doesn’t eat meat.

In theory, reply (a) does not provide a direct answer to (1), and produces a higher processing effort for (a) than does a more straightforward answer like (b). The explanation for the choice of a more costly answer such as (a) is that it provides additional interest (cognitive effects) that could not be obtained from (b) (in this case the reason for her refusal to eat meat), and this interest makes up for the increased effort. Consider now the following ad (cf. Tanaka 1994):

(17) Less bread. No jam (ad by London Transport).

According to RT, hearers follow two basic steps during interpretation: (a) they consider interpretations in the order of accessibility; and (b) stop when their

\(^8\) Indeed, information can be relevant in one context and not in another, so the basic notion they want to define is that of relevance in a context. By a “context” they mean information (“a set of assumptions” in RT terminology) used in interpreting (or “processing”) a new piece of information, either verbal or nonverbal.
expectation of relevance is satisfied. In the case of the ad in (17), the first accessible interpretation is “London Transport is offering something that involves less bread and no jam, probably some type of food”. This interpretation is not relevant in this context (actually, the reader is bound to be puzzled by the combination of “food” and “transportation”). The reader will then consider a second possible interpretation of the ad: “less bread” colloquially means “less money”; hence “no jam” refers to “no traffic jams”. London Transport offers a service which costs less and involves no traffic jams. This second interpretation is relevant in the context of the processing of the ad. Moving to the second interpretation has meant more mental effort for the reader, who is nevertheless satisfied with finding the interpretation. This satisfaction compensates for the effort.

27.8 The explicit/implicit distinction
(in utterance interpretation)

The general two-step procedure for the interpretation of utterances mentioned above is, in reality, a complex cognitive procedure involving a mutual parallel adjustment of three sources of information: (a) the explicit interpretation of the speaker’s utterance (it has to be enriched in order to obtain a fully contextualized proposition); (b) the speaker’s implicated interpretation—implicature—if intended; and (c) the right amount of contextual information needed to get (a) and (b). Unlike Grice’s two-step model of communication according to which one first interprets the utterance literally, concludes that this interpretation is not possible, and then moves on to the implicit or implicated interpretation, S&W predict a dynamic and flexible human cognition capable of accessing context, enriching the utterance at the explicit level, and deriving implicated conclusions without a fixed order, and only constrained by our inherent search for relevance. The parallel subtasks for interpretation are summarized in (18):

a. Construct appropriate hypotheses about explicit content (explicatures) via disambiguation, reference assignment, and other pragmatic enrichment processes (see below).

b. Construct appropriate hypotheses about the intended contextual assumptions (implicated premises).

c. Construct appropriate hypotheses about the intended contextual implications (implicated conclusions).

As pointed out above, interpretation starts with the identification, by the language module, of the schematic and context-free logical form of the utterance which has
to be enriched inferentially. This logical form is turned into a fully contextualized proposition called explicature. To turn the logical form into an explicature (sub-task (18a)) some inferential operations have to be performed (depending on the inferential requirements of the utterance):

— *Reference assignment* and *free enrichment*. Sometimes a reference has to be found for certain words in the utterance. This is typically the case of utterances containing indexicals (i.e., pronouns, adverbs, etc.). Free enrichment, on the other hand, is the inferential completion of the propositional content of the utterance which, despite being apparently complete, needs extra information (e.g., unarticulated constituents) to make sense, as in the bracketed additions in these examples (provided by Carston 2001):

(19) a. Paracetamol is better. [than what?]
    b. It’s the same. [as what?]
    c. He is too young. [for what?]
    d. It’s hot enough. [for what?]
    e. I like Sally’s shoes. [shoes in what relation to Sally?]

This inferential “completion” process is obligatory if the hearer wants to make sense of the intended interpretation and without it there would be no relevant propositional form or explicature.

— *Disambiguation*. When the utterance contains a polysemous word, one of its senses has to be selected according to contextual constraints.

— *Conceptual adjustment*. This is one of the most interesting lines of research within RT. During interpretation, the concept coded by a word is adjusted by the hearers so that it meets their expectations of relevance. The outcome is an *ad hoc concept* which is similar but not identical to the stabilized concept coded by the word.

In certain contexts, the concept that the speaker intends to communicate is *broader* (less exact) than the concept that the word he or she has chosen literally communicates, as in (20a–e):

(20) a. There is a rectangle of lawn in the shed.
    [not an exact rectangle]

Relevance theory’s explicature/implicature distinction is as follows: An assumption communicated by an utterance U is explicit if and only if it is a development of a logical form encoded by U. On the analogy of implicature, S&W (1986: 182) call an explicitly communicated assumption an explicature. Any assumption communicated, but not explicitly so, is implicitly communicated: it is an implicature. Besides explicatures, there are also higher-level explicatures, which include the speaker’s attitude (to regret that . . . , to be happy that . . . , etc.) or a higher-order speech-act schema (to be asking that . . . , to be ordering that . . . , etc.).

They are ad hoc “because they are not linguistically given, but are constructed online in response to specific expectations of relevance raised in specific contexts. There is a difference then between ad hoc concepts, accessed by a spontaneous process of pragmatic inference, and lexicalized concepts, which are context-invariant” (Carston 2002: 322).
b. We entered a pub, but we left since it was *empty*.
   [there were people in the pub—e.g., the waiter—but not interesting people]
c. I’ve got *a thousand things* to do this morning.
   [many things, but not a thousand]
d. Don’t worry. I’ll be ready *in two minutes*.
   [in a while, surely longer than two minutes]
e. This steak is *raw*.
   [not literally raw, but undercooked]

On other occasions, the concept that the speaker intends to communicate is *narrower* (more exact) than the concept that the word he has chosen literally communicates, as in (21):

(21)  
   a. I’ve got *nothing* to wear for the party.
       [nothing appropriate, nothing classy, etc.]
   b. María has *a brain*.
       [not simply a brain: an outstanding brain; she is very intelligent]
   c. This boy has *a temperature*.
       [a higher temperature than he should have]
   d. It will take *some time* to fix this car.
       [longer than you imagine; longer than it would normally take]
   e. Antonio *drinks* too much.
       [drinks too much alcohol]

The notion of *ad hoc concept* is particularly interesting for the analysis of metaphors. The relevance-theoretic account of metaphors is based on the assumption that there is an *interpretive resemblance* between concepts. There is a difference, though, between the initial RT approach and the current one. Indeed, although both accounts rely on the notion of “interpretive resemblance”, in the initial approach the relation was between the concept (or thought) of the speaker and the propositional form of the utterance, and in the new account it is between an encoded concept and a concept communicated.

Besides, both accounts involve the derivation of a range of (strong and/or weak) implicatures. The only major difference is that in the second an ad hoc concept occurs in the explicature, thereby giving inferential warrant to the implicatures derived. In other words, it is claimed that the metaphor provides a new *ad hoc* concept for the proposition expressed by the utterance (which *is* communicated.

---

11 According to Carston and Powell (2005: 283), “while most other pragmatic approaches assume that narrowing and broadening are to be treated as distinct processes, the RT view is that they are simply different possible outcomes of a single pragmatic process which fine-tunes the interpretation of virtually every word”.

12 Implicatures can be stronger or weaker depending on the amount of contextual assumptions that the hearer needs to retrieve in order to obtain them.
as an explication). This concept is part of the information that the hearer uses to derive metaphoric implicatures.

Typically, metaphors involve both broadening and narrowing of concepts, as in utterance (22a) with the intended interpretation (22b), which demands both broadening (22c) and narrowing (22d) (example from Wilson and Carston 2006):

(22)  

a. My daughter is a princess.

b. My daughter, who is not a female royal, is a spoiled, overindulged girl, who constantly asks for special treatment, expects her wishes to be granted, refuses to do housework, etc.

c. The ad hoc concept is broader than the encoded concept in some respects since it applies to some people who are not actual princesses.

d. The ad hoc concept is also narrower in some respects since it applies only to people—including princesses— who are spoiled, overindulged, etc.

In the aforementioned mutual parallel adjustment of explicit content, implicit import, and access to contextual information, the inferential tasks of reference assignment, disambiguation, etc. are applied, when necessary, to the logical form in order to develop the proposition which is communicated explicitly (explication) but, at the same time, this proposition is combined with the right amount of context to yield, again if necessary, fully inferential implicated premises and implicated conclusions (implicatures), and all of these inferential tasks are guided by our biologically rooted search for relevance. Let’s exemplify this with the exchange in (23):

(23)  

Tom: So... Did you buy that table I told you about?

Ann: It’s too wide and uneven.

If Tom wants to understand Ann correctly he has to use inference in order to develop the schematic logical form provided by Ann’s utterance into a relevant interpretation. Some inference will be devoted to obtaining the contextualized propositional form of the utterance which is communicated as an explication. In this particular case, Tom has to engage in reference assignment (“it” refers to “the table”), disambiguation (a table can be “uneven” in several ways: because its surface is uneven or because its legs are not properly leveled), and free enrichment (e.g., too wide [for what?]). The outcome could perhaps be the proposition in (24):

(24)  

Explication: “The table that you told me about is too wide to go through the bedroom door and its surface is uneven.”

This is not the actual answer to Tom’s question, so Tom also has to combine (24) with contextual information (implicated premises) in order to get the intended interpretation (implicated conclusion). In this case encyclopedic contextual information will be accessed by Tom about how unlikely it is for a person to buy a table that
does not go through the door and whose surface is uneven. This contextual information will help Tom reach, as an implicature, the intended interpretation (25):

(25) Implicature: “I didn’t buy the table that you told me about” (*implicated conclusion*).

### 27.9 Relevance and Grammar

Most of the studies of grammar which take RT as the theoretical framework move beyond the traditional view of grammar to a more dynamic and inference-centered approach in which grammatical senses are not taken for granted but supported or refuted according to contextual constraints. In short, grammatical aspects are no longer intrinsic and stable features of language, nor are grammatical attributes a mere list of choices in hypothetical contexts supplied by the grammarian. Instead, a pragmatic and consequently context-centered view of grammar is proposed in which grammatical attributes constrain (or not) the choice of a right (i.e., intended) interpretation. The addressee’s ability to access the adequate context in which the utterance can be optimally processed also plays an important part in the outcome of interpretation. In this case, the grammatical organization of utterances has an important role throughout this cognitive contextualization, since it often imposes constraints upon the range of possible interpretations of the utterance and thus reduces (or increases) the effort required to select the intended interpretation (Yus 1997: 237). A short review of studies which, one way or another, analyze grammatical aspects of language using RT is provided below.

(a) *The conceptual/procedural distinction*. This is one of the most important contributions of RT to the study of grammar, which has resulted from the work by Blakemore (e.g., 1987; 2002) on connectives and discourse markers and has led to a great number of studies in the same area. Instead of a typical approach to connectives, which tends to make a basic distinction between a same-level relationship of elements (coordination, parataxis) and a hierarchy-based one (subordination, hypotaxis), connectives including *after all*, *so*, *but*, *whereas*, etc. are regarded as constraints on relevance, that is, as guidelines for the correct comprehension of the compound sentence, since they reduce the effort needed to access the correct interpretation: “[T]heir sole function is to guide the interpretation process by specifying certain properties of context and contextual effects. In a relevance-based framework, where the aim is to minimise processing costs, the use of such expressions is to be expected” (Blakemore 1987: 77). This minimization of effort can take different directions, since these terms aid in obtaining the speaker’s intended effects by restricting the construction of either the explicatures or the implicatures of the utterance.
Blakemore (2002: 82) lists three possible attributes that we can expect an expression which encodes procedural meaning to have. Firstly, there is “elusiveness”, in the sense that procedural expressions are hard to paraphrase or translate and their descriptions are usually controversial (for instance, there is a difference between the procedural _but_ and the non-procedural _in contrast_, the latter being easier to paraphrase).

Secondly, procedural discourse markers do not have synonymous VP adverbial counterparts. One of the examples by Blakemore (2002: 84) shows how _in other words_ is used in the same sense in (26) and (27), whereas _well_ is used differently in (28) and (29), which implies that _in other words_ encodes conceptual meaning and _well_ encodes procedural meaning:

(26) _In other words_, you’re banned.

(27) She asked me to try and put it _in other words_.

(28) A: What time should we leave?
    B: _Well_, the train leaves at 11.23.

(29) You haven’t ironed this very _well_.

Thirdly, elements that encode conceptual meaning can be semantically complex whereas elements that encode procedural meaning cannot.

Consequently, from this perspective, connectives such as _so_, _but_, and _after all_ are used in order to make implicit coherence relations explicit, and hence to establish a safe guideline for the interpretation of utterances containing them. For instance, the connective _but_ helps the hearer to infer that the proposition it introduces is relevant as a denial of the expectation or as a contrast regarding the proposition expressed in the first clause.

W&S (1993: 10) argue that “inferential comprehension involves the construction and manipulation of conceptual representations; linguistic decoding feeds inferential comprehension; linguistic constructions might therefore be expected to encode two basic types of information: concepts or conceptual representations on the one hand, and procedures for manipulating them on the other”. Connectives such as _but_ should not be seen as encoding concepts but as procedural devices which _constrain_ the inferential phase by indicating the kind of cognitive process that the hearer should go through (hence reducing the eventual overall effort).

In further research on the conceptual/procedural distinction the list of expressions encoding procedural meaning has been extended to other elements of language, for example punctuation marks (Borochovsky Bar-Aba 2003), tense, mood (see below), and even to nonverbal qualities of communication such as intonation (Escandell-Vidal 1998).

(b) _Conditionals_. In general, there is some discussion on the semantic vs. pragmatic uses of conditionals. Smith and Smith (1988) and other authors claim that the RT framework is useful to clarify this point. They suggest that the behavior of both
factual and counterfactual conditionals can be explained in contextual, relevance-theoretic terms. For example:

(30) a. If you are confident enough, bet your whole salary on that horse.
   b. If I ask you politely, will you post the letter?
   c. If you are hungry, there is a flan in the fridge.

In these sentences the hearer has to recover the propositional form of the sentence (via enrichment of the logical form) and integrate it into a description according to the imperative (*the speaker is telling the speaker to p*, as in (30a)), or an interrogative connotation, as in (30b), and in both cases there is a guarantee of relevance for the speaker and/or hearer. For (30c), Smith and Smith propose the following RT-related explanation:

[...] the antecedent specifies a state of affairs which, as usual, provides a relevant context for the consequent. Given the Principle of Relevance, this in turn forces the listener to make certain additional assumptions: specifically, that he can infer from the guaranteed relevance of the consequent that the flan in the fridge is available for him. Given the easily accessible information that hunger is undesirable, that eating alleviates hunger and that flans are for eating, the force of the whole conditional is accounted for naturally. (Smith and Smith 1988: 335)

(c) Modals and modality. Several authors have addressed modals and modality using a relevance-theoretic approach (e.g., Berbeira Gardón 1996, Groefsema 1995, and Nicolle 1997). They attempt to provide a cognitive explanation for the various senses in which modals can be used in similar contexts. For example, *may* has both epistemic and deontic interpretations of (31a) in (31b) and (31c) respectively:

(31) a. She may do the examination tomorrow.
   b. It is possible that she will do the examination tomorrow.
   c. She is permitted to do the examination tomorrow.

Under RT, modals are considered to have a basic meaning, and the different interpretations which they can acquire are dependent on contextual attributes. In other words, there is a basic propositional meaning which is later enriched to yield a propositional form with a context-related (epistemic or deontic) meaning. Groefsema (1995: 61) goes on to say that “the basic meanings of *can, may, must* and *should* express relations between the proposition expressed by the rest of an utterance containing them and a set of ‘background’ assumptions, while putting constraints on what sets of assumptions are recovered during the interpretation process”.

(d) Adverbs and adverbials. Ifantidou-Trouki (1993), among others, has studied adverbs and adverbials under RT. She deals with four types of adverbials: illocutionary (*frankly, confidentially, honestly*...), attitudinal (*unfortunately, happily, sadly*...), evidential (*evidently, obviously*...), and hearsay (*allegedly, reportedly*). They are usually regarded as indicators of the type of speech act performed with the
utterance. The RT-based analysis proves that in reality these kinds of adverbials are very different from each other in their use. Indeed, not all adverbs have a procedural role in the utterances where they occur, helping the hearer in their processing, but they can also encode conceptual representations.

(e) Mood. Clark (1993), W&S (1988), and Ahern and Leonetti (2004), among others, have studied moods such as subjunctive and imperative. Clark’s concern is with pseudo-imperatives. These include verbs used with a covert conditional meaning, as in (32a–d) and also the kind of imperative sense of let’s in (32e):

(32) a. Wash the car and I’ll buy you an ice-cream.
    b. Leave the house or I’ll call the police.
    c. Come one step closer and I’ll shoot.
    d. Turn on the radio and you’ll hear the news about the murder.
    e. Let’s go to the movies tonight.

Clark claims that RT can explain their grammatical behavior, basically proposing that the semantic content of the utterance is combined with contextual information in order to access its intended interpretation. In this sense, the explanation of the “conditional” interpretation of sentences such as (32a–d) is entirely pragmatic: “[I]n each case the hearer has to make some assumption about how desirable the state of affairs is thought to be and from whose point of view it is thought to be desirable; in making these assumptions he is guided by contextual factors and considerations of optimal relevance” (Clark 1993: 82).

On the other hand, Ahern and Leonetti (2004: 37) argue that in the case of verbal mood, its semantic content contributes to the specification of explicatures, mainly those known as higher-level explicatures (attitude- or speech-act-connoted propositions), in which the speaker’s propositional attitude and communicative intention are represented. This conceptualization of mood fits the aforementioned procedural/conceptual dichotomy, since “the semantic content of the grammatical moods is minimal compared to the range of interpretive effects that the use of one or the other can convey: their stable, unitary semantic content leads to a variety of diverse interpretive effects depending on the context they are used in” (ibid.).

(f) The article. Jucker (1992) unifies within the RT framework the different senses of the definite article that are typically proposed in grammars. Hearers can only make hypotheses about their interlocutors’ assumptions, and successful communication depends on these assumptions being accurate. This, when applied to noun phrases, implies that speakers continuously wonder if their interlocutors will manage to identify the intended referent of the noun phrase or not, with the following sub-assumption: “The referent of the expression which contains a definite article is uniquely identifiable at the particular point of the discourse at which the expression occurs” (ibid., 128), which is part of the meaning of the definite article and therefore underlies most of the uses proposed by the grammar: “[T]he various categorisations of uses of the definite article that have been proposed
by grammarians in essence try to compartmentalise the bases on which speakers assume that their addressees will be able to pick up the correct referent” (ibid., 130).

(g) Tense. Authors such as Smith (1990) and Haegeman (1989) suggest that RT provides an ideal framework to avoid the ambiguities that arise in logical descriptions of tense. Indeed, for RT it is not enough to say that the past tense provides a temporal reference prior to the moment of speaking. Instead, the hearer is expected to “narrow the reference down to some more specific interval, so that the utterance can be constructed as expressing an optimally relevant proposition; that is, so that it can interact with accessible contextual assumptions to give rise to a range of effects” (Smith 1990: 85). Smith goes on to apply RT to the sequentiality of the narrative past in sentences such as the following:

(35) John entered the office. The president walked over to him.

in which the information provided by the first part precedes that provided by the second part. Smith concludes that this is part of “reference assignment” and consequently part of the propositional enrichment to yield an explication of the sentence.

Haegeman’s analysis focuses on the difference between going to and will in an attempt to overcome the intuitive distinction of these auxiliaries in examples such as (34a–b):

(34) a. I will/shall leave next week.
   b. I’m going to leave next week.

Her claim is that at the level of sentence meaning, the meaning of both auxiliaries is equivalent, and that the difference is to be found in the constraints that they impose on their processing in the context of the utterance in which they occur. In her study she concludes that,

*Be going to... imposes a constraint on the processing of the proposition with which it is associated. It signals that this proposition is relevant in a context including at least some present tense propositions, or, in other words, it guarantees a contextual effect if the utterance is processed against a present context. Will, on the other hand, signals that the hearer should extend the immediately accessible (present) context for the processing of the proposition and should process the utterance against future propositions.* (Haegeman 1989: 305)

(h) Aspect. Žegarac (1990; 1993) analyzes aspect from a relevance-theoretic approach. Starting off with a traditional classification of aspect and situation types, he concludes that under RT these traditional labels seem problematic. For example, in the case of stative verbs, which cannot be used in the progressive aspect, many instances are found which are clear exceptions to this rule (1990: 127):

(35) a. Peter is being polite.
   b. Mary is loving the fruit salad.

In this case, the hearer realizes that (35a) and (35b) are not intended to be interpreted literally, and starts looking for assumptions about what the speaker might
have intended to convey (normally in the form of implicatures). Basically, the heart of the matter lies in the quality of the encyclopedic entries for the verbs concerned and the hearer’s accessibility to contexts in which the stative meaning of the verbs is eventually created. Some examples by Žegarac (1990: 129) are particularly illustrative:

(36)  
(a) John doesn’t feel well.  
(b) John isn’t feeling well.

(37)  
(a) The baby resembles her mother.  
(b) The baby is resembling her mother more and more.  
(c) ?? The baby is resembling her mother.

(38)  
(a) Antoinette understands Russian.  
(b) Antoinette is understanding Russian better and better.  
(c) ?? Antoinette is understanding Russian.

For Žegarac, the meaning of the progressive is to be defined in terms of reference to instantiation(s) of the property denoted by the predicate, and the predicates in (36–38) take the progressive. However, “feel well” does so more readily than the predicates “resemble one’s mother” and “understand Russian”, acceptable only in (37b) and (38b), which contain explicit indications of change (“more and more”, “better and better”). The difference between the progressive form of “feel”, on the one hand, and “resemble” and “understand Russian” on the other, would be accounted for by RT not as a difference in the degree of “stativity” or “dynamicness” inherent in the meanings of these verbs but as a difference in the accessibility of contexts which are used to achieve adequate contextual effects.

### 27.10 Social aspects of communication

RT has been criticized for avoiding the social aspects of communication. Recent discussion includes Rajagopalan’s (2005) claim that S&W cannot account for intercultural aspects of communication. S&W (1997: 147) acknowledge that they have concentrated on the inferential activity of the individual, but inferential communication is also essentially social: “Inferential communication is intrinsically social, not just because it is a form of interaction, but also, less trivially, because it exploits and enlarges the scope of basic forms of social cognition. Right or wrong, this is a strong sociological claim.” In W&S (2005: 100) they add: “Although pragmatists generally see communication as both a cognitive and a social process, they do not always devote their efforts equally to developing rich accounts of both the cognitive and the social factors involved. We see this as a difference in interests and
of April 8, 2009, 21:55

research strategies rather than in theoretical commitments. In our own work, we have focused on cognitive factors, but we still assume that a comprehensive picture of communication should integrate both kinds of factors.”

As pointed out in Yus (2006), a proposal by Escandell-Vidal (2004) aims at integrating individual and social issues in terms of principles and norms, respectively, and as part of a dynamic picture of human inference. The mind operates according to principles that are in charge of obtaining fully propositional interpretations from coded stimuli. When dealing with norms, the mind is engaged in both a long-term and a short-term task. The latter analyzes and categorizes incoming information, and the former builds up and updates socially accepted behavior.

27.11 Empirical evidence

A common criticism of RT is that it is highly speculative, predicting without empirical evidence the mental procedures and interpretive steps the human mind goes through in human communication. Obviously, we are dealing with an object of study, the human mind, that is highly complex and still largely unexplained. S&W (2002: 143) acknowledge that in much pragmatic research there is a certain reluctance to get down to experimentation. But recent research on a number of pragmatic issues has shed light onto empirical evidence for RT. For instance, Van der Henst and Sperber (2004) review various experimental tests of the two principles of relevance. Other studies have focused on Wason selection task. The plausibility of a Gricean maxim of truthfulness to explain human communication has also been tested. The research showed that when people ask a stranger the time in the street they get, as a reply, “a time that is either accurate to the minute or rounded to the nearest multiple of five, depending on how useful in the circumstances they think a more accurate answer would be” (W&S 2002b: 598), regardless of whether the people asked have (accurate) digital watches. These rounded answers are not strictly true, but more relevant in that context.

27.12 Relevance and other areas of pragmatic research

RT has been applied to many research areas such as humor (e.g., Yus 2003), media discourses (e.g., Tanaka 1994; Yus 1998b, 2001), literature (e.g., Pilkington 2000),
politeness (e.g., Jary 1998), irony (e.g., W&S 1992), translation (e.g., Gutt 2000), and language teaching (e.g., Nizegorodcew 2007) in addition to grammar. These research areas which take RT as their theoretical framework are clear evidence of the dynamism and impact of this cognitive pragmatics theory of communication.

References


