



HAL
open science

A relevance theoretic approach to relevance

Anne Reboul

► **To cite this version:**

Anne Reboul. A relevance theoretic approach to relevance. Acts of the Relevance Theory Workshop, University of Luton, 8-10 septembre 1998, Sep 1998, pp.45-50. halshs-00003805

HAL Id: halshs-00003805

<https://shs.hal.science/halshs-00003805>

Submitted on 2 Feb 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A relevance theoretic approach to reference

Anne Reboul, LORIA-CNRS

Abstract: This paper outlines a model of what a theory of mental representations (TMR) for reference resolution would be. Such a theory would be a specification of relevance theory on the specific point of referents attribution. Mental representations (MR) are conceived as a kind of hinge between referring expressions on the one hand and objects in the world on the other hand. The introduction defines reference and briefly describes the problem. The 2nd § describes the basics of TMR, while the 3rd § introduces the notion of differentiation as a condition for the construction of a MR. The 4th § proposes the notion of domain of reference (DR) as an alternative to Ariel's notion of *accessibility*. The 5th § shows how events and their consequences can be represented in TMR. The 6th paragraph shows how TMR can be used in reference resolution. The 7th paragraph is a conclusion, in which the proximity and compatibility between TMR outlined here and relevance theory are pointed out.

1. Introduction

As Sperber and Wilson (1986/1995) pointed out, linguistic communication is underdetermined and the logical form of an utterance may be less than completely propositional. Indeed, it is one originality of relevance to have shown that pragmatics has a role to play in the specification of truth-conditions, through the enrichment of logical form. Enrichment of logical form has to do with a quite a number of things, among them reference assignment, i.e. the attribution of referents to referential expressions (RE) or at least to those REs which do refer. REs refer just in case the identification of the referent is necessary for the interpretation of the whole utterance. Sperber and Wilson's book did not describe in detail the process through which referents are attributed to RE. Ariel (1990) offered part of such a detailed account, but her proposal has quite a few defects (see this author 1997), chief among which the fact that she does not try to account for reference assignment but limits her approach to coreference. What is more, though she rightly points out that reference resolution supposes the existence of mental representations (MR), she does not propose any kind of hypothesis on what such mental representations could be. Finally, her account does not take into account the fact that reference implies much more than linguistic processes or sheer accessibility of linguistic expressions. This paper aims to outline what a theory of mental representations, conceived as a specification of relevance theory on reference resolution, could be.

2. Mental representations

As said above, reference resolution supposes much more than linguistic ability, for a very simple reason: though reference can be made and very often is made through language, implying as it does non-linguistic reality, it is not primarily a linguistic relation and cannot, generally, be solved through linguistic analysis. Rather, it is a pragmatic phenomenon, which characteristically implies visual and spatial perception, linguistic interpretation, encyclopaedic knowledge and, presumably, proprioception. Leaving proprioception aside, this strongly suggests that any MR used in reference assignment should gather heterogeneous informations. It seems fair to suggest that whereas concepts in general are generic in that they gather informations which enables the individual to decide whether a given object does or does not belong to a given category, MRs are specific concepts which enables one to determine, among the objects belonging to a given category, which is the specific one being referred to. Thus it may be reasonable to suppose that MRs can be built on the general model of concepts proposed in Sperber and Wilson (1986/1995).

Thus a MR should have the following composition: an **address** (which is both its name and a means of access), a **logical entry** (which indicates what logical relations it has with other mental representations), an **encyclopaedic entry** (which gathers both general informations inherited from the category to which the object belongs and informations specific to the object), a **visual entry**

(which may contain an image of the object inherited from its category and which registers changes in its appearance), a **spatial entry** (which may contain informations about the object's intrinsic orientation - if it has one-, and which keeps tracks both of its spatial situation relative to other objects and of its movements), and a **lexical entry** (which gathers all the REs which have actually been used to refer to it). Note that, just as is the case for concepts, all fields do not have to be filled in a MR. MRs are not linguistic objects and, ideally, there should not be more than one MR for a given object.

Quite a few different operations can be applied to MRs, and these can be triggered by either linguistic, perceptual or inferential means: **creation** (the operation through which a new MR is created), **modification** (the operation which allows the addition of new information to a given MR), **fusion** (the operation which takes two MRs and yields only one, reorganising the information in both, when it is recognised that the two MRs in question correspond to the same object), **duplication** (the operation through which a new MR is built from a previous one, used in the treatment of utterances such as: *Take the president speech, xerox it, put the copy on my desk and the original in the strong box*) and two very important operations to which the next section is dedicated, **grouping** and **extraction**.

3. MRs for groups: differentiation, grouping and extraction

Let us consider the two following examples:

- (1) A man and a woman entered the pub. They sat at a table in front.
- (2) Fred had nine marbles. He dropped them. He only found eight. The last one had rolled under the sofa.

(1) and (2) illustrate two types of plurals: in (1), the plural in the second utterance (*they*) designates a couple of well differentiated individuals; in (2), though I know that the marbles are different, I do not have enough information to distinguish between them, or, in other words, to *differentiate* between them. TMR tries to be as realist as possible relative to what type of MRs can be attributed to human beings and it would not be reasonable to insist that objects which cannot be differentiated should receive distinct MRs. Hence, the suggestion is that the creation of a new MR is only allowed when there is a way of differentiating the object from other objects, by whatever means this can be done.

In (1), the difference between the two individuals, a man and a woman, allows the creation of two new MRs, designated here through their adresses, i.e. [*@man*] and [*@woman*]. The coordination of these individuals through the expression *a man and a woman* triggers the **grouping** of the two MRs, which yields a new MR, [*@man⊕woman*]. This new MR, [*@man⊕woman*], has a logical relation with both original MRs, [*@man*] and [*@woman*]. The relation is a relation of partition. Partition is a standard operation in set theory and has the three following characteristics: it constructs subsets in given set; there is no intersection between the subsets of a given partition of a given set (though there may be different partitions of the same set, but this is a different problem); the nul set cannot be the result of a partition.

In (2), though there is a difference between Fred's nine marbles, we are not given enough information to differentiate among them and thus we are allowed only one MR concerning them, [*@marbles*], at the end of the first utterance. The pronoun *them* in the second utterance is resolved on that MR. The situation becomes more complex at the end of the third utterance, where we are given an information which not only allows but makes it necessary to differentiate among the nine marbles, between the eight marbles which Fred found and the one which he did not. This is where the other operation, **extraction**, comes into its own. From the original MR, [*@marbles*], two new MRs are built, on the basis of the conflicting property |found/~found|. These new MRs, [*@1marble*] and [*@8marbles*], bear a logical relation to the original MR, which, just as in the case of grouping, is a relation of partition. In this case, the MR [*@marbles*] is partitioned to yield two new MRs, [*@1marble*] and [*@8marbles*]. Note that in this case the operation of extraction is triggered through inference.

In both the case of grouping and of extraction, the relation of partition is indicated in the logical entry of all the MRs concerned. Finally, one and the same MR can be the object of several different extractions (partitions) and of several different groupings.

4. Domain of reference

Given MRs and the diverse operations which can be applied to them, and which seem to strike the balance between psychological realism and ontological parsimony, there still seems to be a problem: if MRs are the things accessed in reference resolution and allowing the identification of the object the speaker wanted to refer to, they can only be efficient if there are not too many of them, else the selection of the right one will be a very long and costly process. This, allowing for the fact that Ariel (1990) did not speak of mental representations but merely of antecedents, is more or less the problem which the notion of *accessibility* tries to solve. For reasons which I hope will be obvious below, I do not think it sufficient, and, given the alternative device which I will outline in this section (the construction of domains of reference), I do not think it necessary either.

A domain of reference (DR) is a subset of the set of mental representations for a given individual at a given time. Each new RE is interpreted relative to a DR. Its resolution can be either of five things: it can consist of the identification of a MR which belongs to the DR and the corresponding object is then the referent of the RE; it can consist of the extraction from a MR in the DR, yielding a new MR, the object corresponding to that new MR being the referent; it can consist of a grouping of several MRs in the DR, the object corresponding to the resulting MR being the referent; one of the MRs in the DR can give access to another MR with which it has a logical relation; the RE is not referential and a new MR is created and added to the DR. DRs do not include entire MRs. Rather they include the addresses of the MRs, which are a means of access to the MRs themselves. They have a limited capacity and cannot contain more than five to seven MRs (though they may include as few as one MR). What is more, they are not given, but constructed for the interpretation of each new RE. The REs themselves contribute to the construction of DRs through both the conceptual and the procedural informations which they carry. I use the standard *conceptual/procedural* distinction (see Blakemore 1987). According to us, each RE carries information which is used both in the construction of the DR relative to which it must be interpreted and in the selection of the relevant MR in this DR (or in the creation of a new MR if the RE is not referential, but I will not speak of it here for reasons of space). This information can be either both procedural and conceptual or only procedural, depending on the RE employed: for instance, pronouns do not encode conceptual information. No RE carries only conceptual information. In descriptions (NPs with a head noun), the conceptual information is carried by N and the procedural information is carried by Det. I will only give one example, that of definite descriptions. Informally the instruction carried by a definitive description such as *the black cat*, as far as the construction of the DR is concerned, is: *Build a DR in which there are one and only one MR corresponding to an object which is both black and a cat (or one and only one MR yielding access to a MR corresponding to an object which is both black and a cat) and other MRs corresponding to objects which are not both black and a cat.* The instruction concerning the selection of the “right” MR is: *Select the only MR which either belongs to the DR or is accessed through one MR in the DR which corresponds to an object which is both black and a cat.*

The notion of DR can be useful for the analysis of some quantified sentences and its existence is independantly justified (see below and this author, 1995). It seems to us that it can do all that accessibility can and more than it can.

5. MRs and events

Up until now, I have only spoken of MRs for objects. However the importance of events for reference cannot be ignored, on two counts: one can refer to events through REs and if REs are resolved through MRs, there should be MRs for events; events characteristically produce changes of property in objects and this inevitably opens new ways of referring to them. Thus, events must, in one way or another, be included in TMR. The first argument in favour of including them, i.e. they can be

referred to through REs, militates in favour of MRs for events. MRs are means of access to the world and, as such, should integrate at least all or parts of the conditions of identity of the individuals concerned if they are known (these are what allows for distinguishing between individuals, in this case, between events). Roughly, the conditions of identity of a given event are: its nature (the category to which it belongs), its spatio-temporal circumstances (when and where it occurred), its causes and its consequences, its participants, its place in the sequence of events or the sub-events which are parts of it (Conditions of identity for an object include the category to which it belongs, the properties, past and present, which it has at a given moment, its spatio-temporal locations and the events in which it was a participant). Hence, a MR for an event will have much the same fields as does a MR for an object (excepting the spatial entry): its nature, its participants and its spatio-temporal properties are indicated in its encyclopaedic entry as are its sub-events, if it has any. Its consequences are generally the resulting states of one or more of its participants which it has affected, and these are indicated in the encyclopaedic entries of these participants.

This, however, raises another question. Let us take a look at (3):

- (3) The door was grey. A painter came. He painted it red.

In such a case, the door is the object affected by the event described: it has changed color and this change of state must be registered in the encyclopaedic entry of the MR corresponding to it. This however cannot mean a simple substitution of *|is grey|* by *|is red|*, as it is possible to identify an object through one of its past properties as in *The door that was grey* or *The ex-Prime Minister*. Thus information about past states must be preserved though they must not be represented as current states. This means that much of the information in the encyclopaedic entry, and, at least, much of the information specific to the object, must be temporally ordered.

In order to do this, I borrow a very simple system from Romary (1989), based on two temporal relations: **adjacence**, which obtains between any events, which precede or follow each other, and **inclusion**, which obtains between any events one of which occurs during the same interval of time as the other (adjacence and inclusion, obviously, are transitive relations). This allows at least a partial ordering of states in the encyclopaedic entries of MRs for objects. What is more, it also allows the ordering of sub-events in the MR corresponding to an event (if it has any). In such a case, the sub-events are also represented in MRs, the MRs in question being in logical relation with the MR corresponding to the main event. For instance, in (4), there is one main event, described in (a), and three sub-events, described in (b), (c) and (d):

- (4) (a) Fred went to New York. (b) After a violent tempest, the plane was able to land. (c) The passengers came down. (d) Fred was relieved to have safely arrived.

If [*@e(a)*] is the MR corresponding to the event described in (a), [*@e(b)*] that corresponding to the event described in (b), [*@e(c)*] that corresponding to the event described in (c) and [*@e(d)*] that corresponding to the event described in (d), [*@e(a)*] is partitioned in [*@e(b)*], [*@e(c)*] and [*@e(d)*]. The temporal ordering is indicated through the relation of adjacency between [*@e(b)*] and [*@e(c)*] and between [*@e(c)*] and [*@e(d)*] in the encyclopaedic entry of [*@e(a)*]. Note that this means that sub-events are treated through the operation of extraction.

If, instead of (4), we had been told:

- (5) (a) After a violent tempest, the plane was able to land. (b) The passengers came down. (c) Fred was relieved to have safely arrived.

we would have had to construct three MRs, [*@e(a)*], [*@e(b)*] and [*@e(c)*], but it is not clear where we would have indicated the temporal sequencing. I suggest that when we want to order events which are not described as sub-events though they are clearly ordered temporally, we make a grouping operation over the RMs corresponding to them and indicate their temporal order in the encyclopaedic entry of the MR obtained. For (5), the grouping would yield a new MR [*@E*] which would be partitioned in [*@e(a)*], [*@e(b)*] and [*@e(c)*]. In the encyclopaedic entry of [*@E*], a relation of adjacency would be indicated between [*@e(a)*] and [*@e(b)*] and between [*@e(b)*] and [*@e(c)*]. Thus temporal ordering would crucially depend on the two operations of grouping and extraction.

6. *The efficiency of TMR*

I have claimed above that TMR, as described here, would be more efficient for reference resolution than would Accessibility Theory (AT). I would now like to defend that claim. That TMR has quite a few advantages over AT should be obvious: it is a theory of mental representations, which AT is not; it incorporates a notion of DR which can be useful to deal with, at least some examples of quantified utterances, whereas the notion of accessibility cannot be of any help there; this notion of DR can also account for reference resolution in cases where AT would fail. I will begin with the last of these claims and examine (6):

- (6) Jack is a bruttish lout. He has knocked down Jill, his wife.(...). Violence runs in the family: Jill has beaten her daughter very brutally. The victim was transfered to the hospital and the judge was notified.

We are interested in the interpretation of the RE *the victim*. Let us look at it that way: Ariel, not having a theory of mental representations cannot choose between Jill, her daughter or even Jack as antecedent for the victim (there is nothing in AT to say that Jill and her daughter are victims but that Jack is not). Granted, Jack is farther away (or less accessible) than are Jill or her daughter, but the RE is a definite description and definite descriptions allow for a good distance between the RE and its antecedent (i.e. definite descriptions encode comparatively low accessibility). Thus, there is nothing in AT which would allow for a successful resolution of the RE *the victim*. This, however, does not say in anyway why DR should be better than accessibility. Why could not TMR, minus DR and plus AT, do the trick? Let us suppose that we have MRs (such as those outlined here) and that we select them on the basis of accessibility theory (the accessibility of a given MR, for instance, could be downgraded through time when it has no been used for some time in the resolution of REs). So, let us come back to (6): we now have MR and we can calculate the consequences of both Jack and Jill's acts of violence. Thus we know that Jill is a victim and that her daughter is a victim, but that Jack is not. This nicely reduces the field: however, both Jill and her daughter are equally accessible. How, then, can we choose between them? The answer is simple: we cannot.

What then of TMR? What can it do that AT cannot? Admittedly, one would need a good deal of TMR in AT in order to obtain the rather unsatisfying result indicated above. But the main problem is the very notion of accessibility versus DR. MRs for events are regular MRs. As such, they can come in the DR. The interpretation of an action sentence, such as *Jill has beaten her daughter very brutally*, implies the creation of a new MR (this is in agreement with Davidson (1980), who defends the view that action sentences are tantamount to indefinites) for the event described and the introduction of this new MR in the DR. In this MR, the participants, Jill and her daughter, are indicated in their respective roles (the daughter as the victim and Jill as the agressor), and access is given to the daughter (but not to Jill) *via* the MR corresponding to the event. Thus the resolution device finds the "right" referent, i.e. Jill's daughter. It seems to us that, though TMR, still being in its infancy, cannot solve all the problems of reference resolution, it is a more promising direction than is AT.

Let us come now to a problem raised by some quantified utterances (those for which Evans (1985) proposed the notion of E-type pronoun):

- (7) Here some windows were broken by the explosion; they will be repaired by the end of the week at the latest.

The problem is raised by the presence of a pronoun (*they*), which clearly seems to corefer with a quantified NP (*some windows*) without being bound by it. Evans' solution was to say that such pronouns are E-type pronouns and should be analysed as referring to those objects which satisfy the predicate applied to the quantified NP (*were broken by the explosion*). This solution, however, meets with a number of counter-examples, such as:

- (8) Here no windows were broken by the explosion. They are made of unbreakable glass.

Here the pronoun cannot be interpreted as referring to those objects which satisfy the predicate, as there are no such objects. The interpretation, according to us, should go as follows, supposing that *here* means in the building in front of which we are standing: the DR includes (at least) one MR,

corresponding to the building in question ([@building]). This MR is then the object of an operation of extraction by which a new MR corresponding to the windows of the building, [@windows], is built, this new MR then entering the DR. The quantified utterance can then be interpreted as either producing a new extraction on [@windows], between those which are broken and those which are not (in (7)), or as attributing to all the undifferentiated objects corresponding to [@windows], the common property of [not being broken] (in (8)). Thus the notion of DR both can do things which accessibility cannot do and is independantly useful.

Hence, it seems to us that TMR has great advantages over AT. What is more, as I will show in the next section, it may well be closer to relevance theory.

7. TMR and relevance

In which ways are TMR and relevance similar? After all, I have not proposed until now a single way in which relevance or the principle of relevance could play a role in TMR. On that count, AT has emphasized that the notion of accessibility is a way of accounting for the reduction of processing cost through a grammatical indication of where one should look for the right antecedent. So what relations could there be between TMR and relevance, and what is more, are they compatible?

There are a few obvious similarities:

1. first of all the very notion that utterance interpretation (of which reference resolution is only a part) is something far and above what the linguistic periphic system can yield is common between TMR and relevance, though it is not clear that it really is intrinsic to AT;
2. the composition of MRs is clearly based on that of concepts in relevance;
3. the notion of DR is very near to that of context, in that it is built rather than given, and that it is not strictly incremental;
4. inference has an important role in the operations over MRs.

Thus, there should not be any incompatibility between TMR and relevance. The question, then, is: has relevance any role to play in reference resolution as seen in TMR? The first answer to that question could be that TMR has at least as much a role to play for relevance theory as has AT: the notion of DR is as much a way of reducing processing costs as is the notion of accessibility. However, relevance also has a part to play in reference resolution. Given MRs, the operations over MRs and the notion of DR, relevance must, just as it does in utterance interpretation as a whole, intervene for the choice of which MRs should go in the DR (the procedural instructions in the RE play very much the same role as do connectives, that is, they help but they are not enough) and stop the process of reference resolution when the right MR has been accessed. It does that through the usual considerations on costs and effects. In the specific case of reference resolution, once the DR has been built, they either stop the process once a MR has been selected in the DR, or when another MR has been accessed/built (through grouping or extraction), or when that process has been repeated twice, etc. In other words, some operations on MRs (one should mainly think here of grouping or extraction) can be iterated and the arrest of these operations, when they occur during reference resolution, is due to relevance.

Hence, I think that TMR is a promising extension of relevance theory as far as reference resolution is concerned.

References

- Ariel, M. (1990): *Assessing noun-phrase antecedents*, Londres, Routledge.
Blakemore, D. (1987): *Semantic constraints on relevance*, Oxford, Basil Blackwell.
Davidson, D. (1980) \square *Essays on Action and Events*, Oxford, Clarendon Press.
Evans, G. (1985): "Pronouns", in Evans, G. \square *Collected papers*, Oxford, Oxford University Press.

in *Acts of the Relevance Theory Workshop*, University of Luton, Angleterre, 8-10 septembre 7
1998, 45-50

Romary, L. (1989): *Vers la définition d'un modèle cognitif autour de la représentation du temps dans un système de dialogue homme-machine*, Thèse de Doctorat en Informatique, Nancy 1.

Sperber, D. & Wilson, D. (1986/1995): *Relevance. Communication and cognition*, Oxford, Basil Blackwell.

This author (1995): *Are there E-type pronouns? Evans revisited...*, manuscript.

This author (1997): "What (if anything) is accessibility? A relevance-oriented criticism of Ariel's Accessibility Theory of referring expressions", in Connolly, J.J., Vismans, R., Butler, C.S., & Gatward, R.A. (eds): *Discourse and Pragmatics in Functional Grammar*, Berlin, Mouton de Gruyter.