Discourse Representation Theory:
What it is and where it ought to go.

Prof. Dr. Hans Kamp
University of Texas
Austin, Texas

1. INTRODUCTION

It is my task and goal to explain to you what Discourse Representation Theory (DRT) is, and I find this task not to be that easy.

This is not because the subject is so complicated. The difficulty is rather that, during the comparatively few years that have passed since DRT was first conceived, different people seem to have come to identify different things with the name. Someone who wishes to do justice to all - or even a fair selection - of these varying and sometimes incompatible views cannot reply to the question "What is DRT?" with a single, straight answer.

This is the position in which I find myself today. Yet these very circumstances make the challenge especially welcome to me. The reason for that is twofold. First, I am aware that I have contributed, even wittingly, to some of the reigning confusions over what DRT is, and I would like to straighten some of these out. Secondly, and more importantly, sorting out the different strands of what has been going the name of DRT serves a purpose that extends far beyond the theory itself. For, as we will see, any attempt to understand and separate these different strands must confront certain basic questions concerning the nature of language meaning, its correlation with grammatical form, the relationship between language and thought, and the methodology of linguistic theory; and ultimately they touch on the very possibility of theoretical linguistics as that discipline is conceived by many today. I hope to present some of these issues to you as we go along.

2. DISCOURSE REPRESENTATION THEORY

In order to tackle these loftier aspects of the task, I must first give you an idea of how DRT works.

One of the problems one faces in trying to decide what DRT is is to know how much is to be included. On the one hand there exists a narrow conception, according to which DRT is a specific account of a fairly limited range of phenomena, concerning the semantical and logical properties of certain types of English singular noun phrases. This is a theory which makes definite predictions about the possibilities of pronominal anaphora and about the truth conditions of a certain class of sentences in which the relevant NP's play a prominent part, and which has the scientific respectability of being falsifiable. (I would now say that it has unequivocally passed the test of respectability in that it has in fact been falsified. But more about this anon.) At the other extreme of the spectrum is the notion of DRT as a general approach to the explanation of linguistic meaning, in which the concerns and requirements of modeltheoretic, or "formal", semantics are combined with a more procedural conception of the way in which linguistic meaning relates to linguistic form. Here, we have little more than a methodological mould into which many particular theories, indeed many mutually incompatible theories, can be fitted. As a mould it may have its merits, in so far as it can help us to think about problems in a certain productive way. But it could hardly be *refuted.* It can only fall into disuse, if and when its capacity for inspiring solutions to particular problems will have been exhausted and it no longer serves as a meaningful paradigm.

Between these two extremes there are a number of intermediate conceptions of what DRT might be, and it is on one of these I would like to focus first. According to this particular conception, an essential part of DRT is the account it has to offer of temporal reference. My reasons for starting with a closer look at this particular aspect of the theory are partly historical. For it was in fact in the context of the study of tense and aspect that DRT, or at any rate my end of it, began.

3. TENSE AND ASPECT

It was in the summer of 1978 - the first of a long series of always pleasant and often productive periods that I have had the fortune to spend at the University of Stuttgart - that I was con-

fronted with the problem of how to account for the difference between the continuous and the non-continuous past tenses of French, the Imparfait and the Passé Simple (PS). This is a notorious problem of French Grammar, and linguists of various theoretical persuasions have tried their hand at it. The interest it has aroused derives partly from the notorious difficulty which for instance Germans and Scandinavians, in whose own languages the distinction between PS and Imparfait is not realized, have in grasping the correct use of these tenses. As a consequence, teachers of French have long been eager for a clear account of the functions of PS and Imparfait and for effective recipies which can tell us how they ought to be used. However, effective recipies have proved surprisingly elusive.

Despite the interest that the problem has aroused, there existed, at the time when I encountered it, no account of the difference between PS and Imparfait that deserved to be called a theory. There were numerous detailed, and often highly sensitive and insightful descriptions of the meaning or function of particular occurrences of PS and Imparfait in particular contexts; and there were a few descriptive generalizations, such as for instance that

(C1) The PS often conveys punctuality whereas the Imparfait conveys temporal extendedness.

and

(C2) The PS has the capacity of driving the action of a narrative forward, but this is a power which the Imparfait lacks.

I have singled out these two from a slightly larger list of contrasts that can be found in the literature, partly because they were the two generalizations by which I was most preoccupied at the time, and partly because they form a rather curious pair (which was one of the reasons why they struck me). On the face of it (C2) looks like a very different sort of principle from (C1), and one of the intriguing questions concerning PS and Imparfait is how two such apparently incommensurable contrasts could be associated with the same pair of linguistic items.

In 1978, the model-theoretic approach to problems of temporal reference in natural language was still a relative novelty, and as the method had already led to solutions of some of those problems, it seemed natural to try it on the present one as well. However, as became soon evident, the existent model-theoretic approach towards the semantics of time was woefully unsuited for this particular problem.

In a nutshell the difficulty is this: Model-theoretic semantics tries to give an account of meaning in natural language by articulating the conditions under which the sentences of a given language are true. But since truth has to do with how things are - whether a given sentence is true depends typically not just on what it means but also on what the subject matter of which it speaks is actually like - the articulation of truth conditions must be relative to some independent reality to which the sentence is, so to speak, *accountable*. In model theory this independent reality enters into the characterization of truth conditions in the form of so-called *models*. A model is a structure which, in some supposedly objective and non-question begging fashion, contains the kind of data which determine the truth or falsity of the sentences of the given language. That the model encodes these data in a direct and natural way is important. For only if we can recognize models as capturing the facts of a given possible world or situation in much the way in which that information is accessible to us as speakers will it be possible to compare what the theory has to say about truth with the judgements which speakers of the language make about the truth or falsity of those sentences in individual circumstances.

In particular, when we articulate truth conditions for a tensed language, we want to use models which incorporate a faithful representation of real time - much say, as time is claimed to be by physics, i.e. as a structure isomorphic to the real numbers. In such a model, an event will count as punctual if it "lasts" for just one instant, which means that it has no real duration at all. But evidently the events that we refer to in ordinary writing or speech - and more particularly, those to which French speakers refer in the PS - are rarely if ever punctual in this extreme sense. If there is any truth in claiming that, in a home and garden variety example such as:

1. Jean écrivit une lettre a sa mère.

the PS conveys an element of punctuality, it surely cannot be punctuality in that sense, for we all know perfectly well that an event such as writing a letter isn't over that quickly.

Clearly then, the punctuality spoken of in (1) must be punctuality in some other sense. In fact, it doesn't require much reflection to see what sort of punctuality must be involved: PS events are "punctual" in the sense of playing a *punctual role* in the story in which they figure. They are, as one might put it, punctual not in the sense of real but of *narrative time*. This view of the matter is also supported by another fact, viz. that there are many events to which we can (and must !) refer with the Imparfait in one context and with the Passé Simple in another. This, as some other phenomena relating to the difference between PS and Imparfait, can be replicated in English, where the opposition between the simple past and the past progressive of non-stative

verbs fairly closely parallels that between PS and Imparfait in French. By way of illustration consider the following three passages:

2. Susan's alarm clock rang at 7:30. She turned it off. Then, after 10 minutes, she got up, made her way to the kitchen, switched on the electric kettle and lit the burner underneath the milk pan. Next she went to the bathroom and had a shower.

3. Alan was just squirting the shaving cream onto his face when the phone rang. Susan was having a shower. So he rushed to the kitchen and picked up the receiver.

4. It was a tense moment. The phone was ringing, the baby was yelling, the kettle was screaming; at least the milk, which had boiled over and now covered most of the stove, did not contribute to the noise. Susan was having a shower.

Comparing (2) with (3) we see that an event which the former presents in the simple past, viz. the event of Susan taking a shower, is referred to in (3) with the help of the past progressive. Similarly, the simple past the phone rang of (3) corresponds to the past progressive the phone was ringing in (4). Intuitively, it seems to be clear what it is that permits the switch from the simple past took a shower in (2) to the progressive was taking a shower in (3), and again that from rang in (3) to was ringing in (4). These switches are a reflection of the fact that the three passages have different *narrative time structures*. The time of (2) is of a "coarser grain" than that of (3), and that of (4) is grained more finely even than that of (3).

If the punctuality - non-punctuality contrast between PS and Imparfait has to do with narrative time rather than real time, its explication will require a theory of narrative time. Indeed, from what I have just said, it would appear that what should be explained is how the two tenses help shape the narrative time structures that are conveyed by the passages in which they occur. Each of the PS and the Imparfait, we may assume, imposes its own constraints on the temporal structure of narrative time, and it seems natural to think of these constraints as rules, which the recipient of a discourse, whose understanding of the story told must include a grasp of its temporal structure, applies when constructing an interpretation of the discourse.

To attempt an account of PS and Imparfait along these lines may seem natural enough. But to turn the idea into something that deserves to be called a theory is another matter. In fact, if we take the present line of thought really seriously, we find that we have taken upon ourselves

a task of major proportions. First, if we are to state the rules associated with PS and Imparfait with any degree of precision, we must specify what sort of structure narrative time in general possesses as it is construed by the interpreter of a discourse or text; and we must say what punctuality comes to in the context of such temporal structure. For only then can we hope to say something non-trivial about the way in which the interpretation rules for PS and Imparfait contribute to narrative structure, and to show that the rules by which they are interpreted do indeed confer upon the contributions they make a punctual c.q. non-punctual status. Second, if the explanation is to be of any real value, then it must be shown in what ways not just *these* two tenses make their specific contributions to the constructed interpretation, but also what contributions are made by other linguistic items which contribute information of a temporal sort - i.e. the remaining tenses, together with the varied collection of temporal adverbs, prepositions and subordinate clause constructions that are found in a language like French. In other words, the particular interpretation rules for PS and Imparfait must be presented as part of a *system* of rules accounting for the full spectrum of devices by means of which the language refers to time or temporal relations.

Once we have gone this far, however, there is no good reason why we should stop here. If it is along these lines that we are to account for the meaning of the tenses and other linguistic devices of temporal reference, what justification could there be for dealing with the semantics of non-temporal aspects of meaning in a radically different way? After all, a human interpreter has to attend to all sentence constituents of a sentence, non-temporal as well as temporal, in order to arrive at a full and correct interpretation of any given sentence. While this consideration does not prove anything, it nevertheless seems to suggest that we have here a method of explanation which should not just work for some elements of a language, but for all. It is true, that the temporal system of a language such as French enjoys a special and partly autonomous status within the grammar as a whole, and it is perhaps conceivable that a special kind of meaning follows from that special status. But this is a position that should be adopted as a last resort. As long as we have no pressing reason for adopting it, we should aim for a natural language semantics that is built entirely along the lines I indicated.

The particular proposal I put forward for explaining one of the contrasts between PS and Imparfait thus leads us towards a blueprint for semantic theory in general, and it is one which differs significantly from any of the compositional theories of meaning that were current in the days when DRT was born. The blueprint is quite different, in particular, from the model-

theoretic approach towards natural language semantics known, after its creator Richard Montague, as Montague Grammar (Montague 1970). (Nevertheless, there are, as I will explain below, also important points of agreement between Montague Grammar and DRT - so much that the latter can be seen as a natural descendant from the former.)

Let us, in the light of these general considerations about the form which an explanation of (C1) might take, have a brief look at the second of the two mentioned contrasts, (C2). As you will remember, this contrast concerns the power which the PS has, but the Imparfait lacks, to move the action of a narrative forward. Perhaps it is already less of a mystery than it had seemed intially how the same principles might be able to account both for this contrast and for (C1), for the processing rules of which we have spoken will have to determine - among other things - how the information carried by a sentence in PS or Imparfait is "slotted into" the representation obtained from the antecedent discourse. And the power to advance narrative time is clearly a matter of precisely how the information carried by the new sentence is connected with that part of the narrative's representation that is already in place when this sentence is interpreted.

4. SINGULAR NOUN PHRASES

We will see eventually in more detail how the right rules for PS and Imparfait can account for both (C1) and (C2). But first I will say something about another application of DRT to a comparatively narrowly circumscribed set of linguistic facts. This is the theory propounded in Kamp 1984. I obliquely referred to it earlier as the most narrow conception of DRT. I also mentioned that it concerns the semantics of certain (English) singular noun phrases. Such noun phrases have been the subject of logical and linguistic investigations for many years, and there exist several explicit treatments of the referential and quantificational aspects of such noun phrases. The possibility of comparing the new theory with those earlier competitors has undoubtedly helped to raise a certain interest in it in the early eighties. In some instances, however, it seems to have fostered the view that the merits of DRT reside wholly in its ability to deal with a small number of data which the extant accounts had left unexplained. These data relate to what is now generally known as the 'problem of the donkey sentences'. The problem is exemplified by sentences such as

5. If Pedro owns a donkey then he beats it.

and

6. Every farmer who owns a donkey beats it.

It would take us too long to go into the question precisely why such sentences have posed a problem for older semantical accounts of reference and quantification. Roughly, the problem is of how to explain how the indefinite description a donkey in the antecedent of (5) and in the relative clause of (6) can "bind" the pronoun it. If we think of indefinite descriptions as devices of existential quantification, as older theories explicitly or implicitly assumed, then one encounters an apparently unresolvable conflict between, on the one hand, the requirement, imposed by the anaphoric link between indefinite description and pronoun, that all the other semantic material contained in these sentences be in the scope of this quantifier, and on the other hand, the requirement (entailed by the meaning the sentences are perceived to have) that the quantifier should not have maximal scope: In (5) it should be confined to the antecedent and in (6) it should be inside the scope of the universal quantifier expressed by every farmer.

This difficulty disappears in DRT because noun phrases beginning with a or some are not treated as existential quantifiers. I should emphasize right away that this was not an ad hoc decision made for the explicit purpose of overcoming the donkey sentence problem, but a natural consequence of the procedural, interpretative approach our blueprint dictates.

5. DISCOURSE REPRESENTATION STRUCTURES

Although the story DRT has to tell about the donkey problem has been recited often enough, I will give a brief recitation of it here once again, for it is after all a good way of bringing to the fore how the theory works and what some of its central assumptions are.

One of those assumptions is that the processing of simple predications such as

7. Anke irritates Helmut.

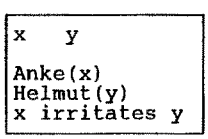
involves setting up representatives for the individuals that the predicating sentence refers to. The representational elements which play this role are called *discourse referents, or drefs* for short.[1] The representation of a sentence such as (7) will contain drefs corresponding to the proper names Anke and Helmut. In addition it must contain certain so-called *conditions*, which capture the descriptive information carried by the sentence. Typically this information is one of two kinds:

(i) information about the identity of the individual which a given dref represents - for instance that the dref introduced for the individual denoted by the name Anke (assume that this dref is x) represents someone to whom it is possible to refer, in the given context, with the name Anke. The condition representing this information has the form:

8. Anke (x)[2]

(ii) Secondly, there is the information contained in the actual predication which the sentence makes, in the sense in which for instance (7) predicates the verb irritates of the individuals Anke and Helmut. In current DRT the conditions that represent this kind of information are obtained by replacing the argument NPs of the predicating clauses by the corresponding drefs. Thus (7) leads to a structure consisting of two drefs (say x and y) and three conditions. Presented in the commonly used diagrammatic form the structure looks like this:

9.



---

[1] Discourse referents are also called reference markers. This is a regrettable duplication of terminology, which I am personally to be blamed for. I am afraid that it is probably too late to do anything about the matter now, and even today my own use is not fully consistent. On the whole, I now prefer the earlier term discourse referent to the later alternative.

[2] The question what proper names refer to and what descriptive content, if any, they carry has been extensively debated ever since Kripke's Naming and Necessity (Kripke 1972). Indeed, much more should be said about conditions like (8) than is possible here. This is an issue about which DRT has some things to say which seem to me to be both novel and important. See e.g. Kamp 1986.

Structures such as (9), which DRT postulates as representations of sentences and sentence sequences, are called *Discourse Representation Structures*, or, more succinctly, *DRSs*.

The description I have given of (9) does not make fully explicit by what algorithmic process it is constructed from the sentence it represents. A precise articulation of this process is essential to the theory, but it is a matter that involves too many details to permit adequate explanation here. A rigorous articulation presupposes in particular an exact specification of the syntax of the language or language fragment for which the algorithm is to be defined. This syntax will assign *a syntactic structure* to each sentence; it is this structure which acts as input to the construction algorithm. In the version of the theory I have in mind here, the algorithm has for each of a range of semantically significant syntactic constituents a corresponding construction rule, and the processing of a syntactic structure consists of applying the rules associated with the relevant constituents of that structure in an order dictated by the structure's hierarchical organization. We will have more to say about particular construction rules as we go along. In the case of (7) the only construction rule used is that associated with proper names. This rule provides for

i) the introduction of a new dref;

ii) the introduction of a condition saying that the dref stands for a bearer of the name; and

iii) substitution of the new dref in the argument position previously occupied by the name.

In the case of (7) this rule has to be applied twice, once to the grammatical subject Anke and once to the direct object Helmut. It is easy to see that the result will be a structure of the form (9).

6. ANAPHORIC REFERENCE

It is common and natural linguistic practice to continue after a sentence like (7) with a sentence which refers back to the individuals that were introduced in the preceding sentence. Among the various types of noun phrases that languages such as English and French have available for the purpose of such "anaphoric" reference, there are in particular the third person singular pronouns. Thus (7) might be followed by

10. He mystifies her.

The natural interpretation of (10) in a context where it directly follows (7) is that he refers to Helmut and she to Anke. A decent theory of semantic interpretation should be able to account for these anaphoric references. In particular, the present theory must specify processing rules for the pronouns which allow for such anaphoric connections in precisely those situations in which they are in fact possible. Thus the rule for the pronoun he, for instance, must specify that the representative for its denotation, i.e. the new dref which the pronoun introduces, is linked with some dref that is already present in the representation. This "antecedent" dref, must satisfy a number of constraints. In the present case, where the pronoun is the masculine he, one of these constraints is that the antecedent dref can be understood as representing a male person. (Normally a number of other constraints have to be satisfied as well, but this is another matter into which I cannot go here.) If we apply this rule to the two pronouns of (10) in the context established by the already constructed DRS (9), we obtain:

11.

x	y	u	v
Anke(x)			
Helmut(y)			
x irritates y			
u = y			
v = x			
u mystifies v			

(Here we have chosen u and v as new drefs for he and her and have used the more or less standard practice of representing the linking conditions as equalities.)

7. INDEFINITE DESCRIPTIONS

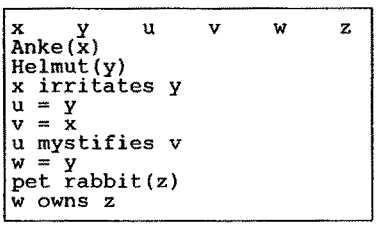
Indefinite descriptions, which, as we saw, play a central part in the donkey sentence problem, require yet another kind of processing rule. This rule shares the feature, that it always introduces a new dref and substitutes this dref for the NP in the argument slot which the NP occupies, with all others that are associated with NPs, and in particular with the rules for proper names and pronouns which we have already met. Once again, the difference is the descriptive condition deriving from the NP itself. In the case of an indefinite description this condition must say that the individual represented by the dref satisfies the common noun phrase following the

indefinite article a. The condition representing this information consists of the common noun phrase followed by the dref in parentheses. Thus, if we continue the mini-discourse consisting of (7) and (10) with the sentence

12. He owns a pet rabbit.

the rules for pronouns and for indefinite descriptions will yield the following extension of (11):

13.

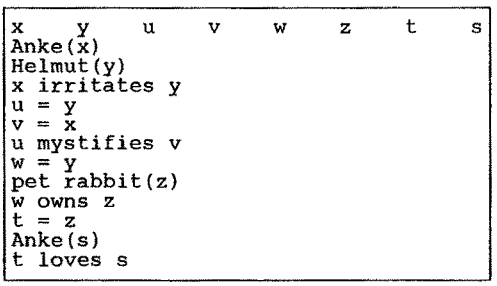


Indefinite descriptions too can act as anaphoric antecedents for subsequent pronouns. Thus we could continue our little discourse with

14. It loves Anke.

meaning by it the mentioned pet rabbit. The DR-theoretic treatment of pronominal anaphora predicts this, for it requires only the presence of a suitable antecedent dref, and the one introduced for the NP a pet rabbit of (12) qualifies. (Whether that dref was introduced by processing a proper name, an indefinite description, or some other type of NP has ceased to be an issue at the point where it is called upon to serve in this anaphoric capacity !). So, applying the rule for pronouns and the one for proper names to (14) we obtain the further expanded DRS:

15.



8. TRUTH DEFINITION IN DRT

This is a good point for saying what DRT has in common with Montague Grammar. Earlier I asserted that the central idea behind model-theoretic semantics in general, and behind Montague Grammar in particular, is that one should account for the semantics of a given language by specifying the truth conditions of each of its sentences. This is a notion to which DRT also subscribes (even though it also embodies the view that there is more to a theory of meaning than truth conditions alone). But in DRT the truth definition applies not directly to sentences of the language in question, but rather to the DRSs that can be constructed from them. The truth conditions of sentences are defined only indirectly viz. via the truth conditions of the DRSs that can be constructed from them, and the same applies to larger discourse units, consisting of two or more sentences. (This is important because - this is another important motive behind DRT - it often happens that a sequence of sentences cannot be analyzed as a simple conjunction of the sentences that make it up. Compare, for instance, the combination of (12) and (14) which, to put it in traditional logical terms, are jointly equivalent not to a conjunction but rather to an *existential quantification* over a conjunction.)

The details of any definition of truth conditions must clearly depend on the kind of the semantic material that may be contained in the objects (sentences, DRSs or whatever) to which the definition is meant to apply. In DRT this is a matter of the semantic material that can be found in the DRSs; but this is in turn determined by the language or language fragment from which those DRSs can be constructed. In particular, the language will determine what sort of structures can serve as models. Different languages will be sensitive to different aspects of reality; so a structure that contains all that is relevant to one language, may nevertheless fail to incorporate some that is relevant to another. So it may be usable as a model for the first language but not as a model for the second. The particular version of DRT we are at present discussing is concerned with a quite simple fragment of English, for which the comparatively simple models of extensional predicate logic turn out to be adequate. For more extended fragments, involving plurals or intentionality, this will as a rule no longer be so. And the more complex model structures that are needed in relation to those richer fragments often entail substantial modifications in the truth definition itself.

But while the details of the definition will vary between one DRT-application and the next, there is one central idea which all those definitions share. In connection with simple DRSs such as (9), (11), (13), and (15) this idea can easily be explained. Note that, formally speaking, each

of these DRSs is a kind of partial model, one which does not necessarily list all the individuals there are and need not decide every possible predication involving those that it does list. Such a structure should count as true, relative to a real model M, if it presents a correct picture of part of M, i.e. if it is possible to correlate individuals of M with the drefs of the DRS such that all the conditions of the DRS are in fact true in M of the individuals correlated with the drefs occurring in those conditions. (For those who like slogans: In DRT truth = correct embeddability.)

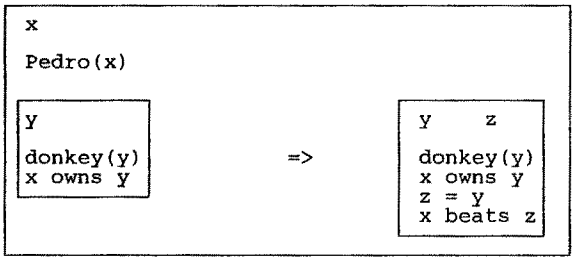
It can be verified that if truth is defined along these lines for the DRS's we have constructed then, in each case, the passage which the DRS represents does indeed get assigned the truth conditions which we take them to have.

9. CONDITIONAL STRUCTURES

Sentences like (5) or (6), which involve such logical concepts as conditionalization and universal quantification, cannot be represented by the simple kind of DRS which we have been led to construct so far. For instance, there is no way of representing the conditional character of (5) with the help of nothing more than the simple conditions which can be found in (9) - (15). Rather, the DRS for (5) must contain some conditional structure which reflects the hypothetical character of the represented sentence. The structure that DRT has adopted to this end plays the role of a DRS condition, while its constituent parts, corresponding to antecedent and consequent of the represented conditional, are themselves DRSs. These subordinate DRSs are constructed from the antecedent and consequent of (5), by the same methods that we used earlier to obtain the DRSs (9) - (15). Moreover - and this asumption is crucial to the solution DRT offers for the donkey sentence problem - the "consequent" part of this hypothetical condition must be seen as an *amplification* of its "antecedent" DRS. That is, we interpret the conditional as a claim to the effect that the situation or situations which fit the desription provided by the antecedent DRS also satisfy the amplified description that is obtained by incorporating the consequent of the represented sentence into that DRS. This assumption is crucial to our problem in as much as it implies that the consequent of the represented sentence is to be interpreted in the context of the DRS which represents the antecedent. This means, among other things, that a pronoun occurring in the consequent may "pick up" any suitable dref which occurs in the

antecedent DRS; and thus in particular that the supposedly problematic anaphora of (5) is indeed possible. The DRS we get for (5) by proceeding along these lines looks as follows:

16.



As it stands, this isn't yet a solution to the puzzle, for we still have to spell out the truth conditions for DRSs like (16). Our discussion of the truth conditions of the earlier DRSs (9) - (15) doesn't settle the matter, for none of those had conditions of the complex, hypothetical sort we find in this latest DRS. However, what we have already said about such hypothetical DRS conditions provides at least a rough indication of what they contribute to the truth conditions of the DRSs containing them: The force of a condition of the form:

17. $K_1 \Rightarrow K_2$

where K_2 is a DRS expanding K_1, is that situations of the kind described by K_1 also satisfy the amplified description provided by K_2. This specification is still rather vague, and can be sharpened in more than one way. This vagueness is nothing other than the notorious equivocality which infects almost all natural language conditionals, and which has given rise to the extensive logical literature concerned with the formulation and exploration of alternatives to the material implication of classical propositional logic.

As it turns out, this is an issue orthogonal to the donkey sentence problem, and as far as I can see any of the established alternatives to the material, conditionals will lead to essentially the same solution of that problem which we get when we treat natural language conditionals as material implications.

The theory I am now expounding adopted (for expository reasons rather than out of any deeper conviction) the material account, and I will stick to that here. For DRS conditions of the type (17) the material account comes to this:

(17) is satisfied in a model M if every situation in M satisfying K_1 also satisfies K_2.

This still leaves one notion to be explained: What are the *situations in* M that satisfy some given DRS K? Here again there may be different options; but in relation to the comparatively simple first order models which we have been using, the only natural one that I can think of is that which identifies those situations with those ways of correlating individuals of M with the drefs of K which render all the conditions of K true in M. In other words, situations are identified with correct embeddings. Adopting this identification we arrive at the following clause for the truth conditions of (17):

18. $K_1 \Rightarrow K_2$ is *true in* M *relative to* some embedding f (f should be thought of as an embedding for the DRS containing $K_1 \Rightarrow K_2$ as a condition) iff any way of extending f to a correct embedding g of K_1 can be extended to a correct embedding h of K_2.

(The need to speak of three different embeddings in (18) arises from the circumstance that some of the drefs occurring in the conditions of K_1 may also occur in other parts of the DRS containing $K_1 \Rightarrow K_2$, while drefs in conditions of K_2 may occur in K_1 as well as in other parts of the containing DRS.)

It should not be hard to verify that (18) assigns to (5) truth conditions that are at least approximately right. Moreover, (they are strictly correct or not) they are in fact the truth conditions which, rightly or wrongly, earlier theories tried to secure, but could not get in any natural way.

For the universally quantified sentence (6) DRT offers an analysis that is in all significant aspects like the one we have just sketched. We omit the details.

10. DRS AS A KNOWLEDGE REPRESENTATION LANGUAGE

There are two general comments I want to make about the theory of referential and quantificational NP's parts of which were outlined in sections 5 - 9. One of these can wait until after we have resumed the discussion of tense which we interrupted in section 5. The other I will make now.

The DRSs about which we have been talking can be regarded as formulas of some artificial language. How powerful this language will be depends of course on the expressive resources of the underlying natural language or language fragment from which the DRSs derive. For the

English fragment that defines the scope of the theory we have been discussing, and which, as I have indicated, is quite limited in its resources, the corresponding DRS language has the power of first order predicate logic but no more. In fact, if we want to convert a DRS K of this language into an equivalent formula of standard predicate logic, all we need to do is to apply the relevant clauses of truth definition to K and write out the result in standard logical notation.

While this is an obvious point, it is also one which has, in my opinion, often been misinterpreted. The equivalence between the DRS language outlined in the preceding sections and classical first order logic has been cited - sometimes with disappointment and sometimes with glee - as showing that DRT has nothing new to offer to knowledge representation. Evidently, so the argument goes, as a representation language these DRSs have nothing new to offer, since they are straightforwardly translatable into formulas of a language which has been with us for more than a century.

This view requires two comments. First, while it is true of the particular DRS language under discussion that its powers coincides with those of the first order predicate calculus, this is not so for the DRS languages corresponding to richer natural language fragments including devices of temporal reference, plural NP's, intensional constructions, and so on. Some such more powerful languages have been in existence for some years now, though it is true that few of the papers in which these extensions are described have been easily accessible.

Secondly, even if no DRS language did go beyond first order logic, this would not trivialize DRT from the perspective of knowledge representation. It should be kept in mind that the repesentation of knowledge raises at least two distinct questions. One is that of finding regimented languages in which a particular body of knowledge can be represented in a logically transparent way. The second issue arises only after the first one has been solved. This is the problem of how knowledge encoded in natural language can be systematically recast in the logically more tractable forms offered by the chosen representation language.

This second problem is every bit as serious as the first. Indeed, it might be said that, insofar as predicate logic can be considered the knowledge representation language *par excellence* (as well as *avant la lettre*), it is precisely this second component of a fully explicit theory of knowledge representation that had been missing throughout the nearly eighty years which separated the publication of Montague's English as a Formal Language from the appearance of Frege's Begriffsschrift (Frege 1879). Montague was the first to show - contrary to popular prejudice - that such a component could in fact be articulated for non-trivial parts of English, but his methods do not appear suitable to address the phenomenon of cross-sentential connections and

their impact on the meaning of discourse. Theories that are to provide translation algorithms from natural language to first order logic which take these aspects of discourse meaning into account, must, I am convinced, incorporate some of the central insights on which DRT is based and some of the theoretical features that derive directly from those insights.

The topic of knowledge representation leads naturally to the closely related issue of *deduction.* When I say "closely related" here, I am making an understatement. For in fact it seems to me that questions of knowledge representation and questions of deduction are inseparable. Any format for representing knowledge is otiose unless its representations sustain correct and efficient mechanisms for *using* the knowledge they represent, and that, I believe, always comes down in the end to deriving from certain given representations others which are logically entailed by them.

This means in particular, that DRSs qualify as serious candidates for knowledge representation only if they furnish the basis for a suitable deduction algorithm or, to use the term current in artificial intelligence, a suitable *theorem prover.* For the DRSs of the theory discussed in sections 5 - 9 - which, as we noted, form a notational variant of first order logic - this problem has of course been solved, at least in the abstract. For in a sense we have had theorem provers for first order logic ever since Frege formulated his rules of proof. Moreover, much is now known both about efficient implementations of deduction methods that were originally developed without any direct concern for questions of automation and about deduction algorithms developed for specific implementational ends. All these methods can be straightforwardly adapted to the DRS language under discussion, given that the DRSs of this language are so easily converted into equivalent predicate logic formulas.

These last considerations speak neither against nor in favour of the proposal to use DRSs for knowledge representations. There is, however, another consideration that is relevant here. Those structural features that distinguish DRSs from formulas of standard predicate logic are highly suggestive of certain inference principles which directly exploit those features and which had not previously been used as separate rules or axioms in first order deduction. These principles have a number of intuitively plausible properties, which at one point seemed to justify the hope that on their basis it might be possible to develop theorem provers of a previously unattained efficiency. Unfortunately, these expectations have not so far been confirmed, and so the strongest arguments for DRSs as knowledge representations relate to their systematic connection with natural language rather than their value as vehicles for logical inference.

11. REPRESENTATIONAL AND PROCEDURAL ASPECTS OF DRT

The second comment is the following: As I indicated earlier, DRT has sometimes been identified with the specific but limited theory of sections 5 - 9. But in addition, not everyone who has identified DRT with this particular theory has intended quite the same thing. Some have had a theory in mind, which was identified by the particular predictions the theory makes about the possibilities of pronominal anaphora and the truth conditions of sentences in which pronouns or other types of singular noun phrases occupy an important position, and be prepared to see the representational and procedural aspects of the theory - which I have been stressing here as central to its conception - as having at best heuristic importance. Others have regarded precisely those aspects as essential to the theory and have wanted to see its tangible success (such as it is) in making predictions about anaphoric possibilities and truth conditions as a confirmation of that these aspects are important to natural language semantics.

The difference between these two perspectives moves into sharper focus when the theory I have sketched is compared with Irene Heim's File Change Semantics, henceforth FCS (Heim 1982). FCS and the theory of the preceding sections were developed independently and roughly at the same time (i.e. around 1980). With regard to those linguistic data to which both theories apply, they make the same predictions; moreover, the mechanisms by which they arrive at these predictions are importantly similar as well. In particular, both theories make a crucial use of the notion of context, stipulating for instance that the antecedent of a conditional acts as a context which partly determines the content of the consequent. But Heim's theory makes no reference to representations, nor does it pretend to say anything about the procedural aspects of language interpretation.

Evidently someone who considers FCS and DRT as essentially the same theory must either read representational and procedural elements into the former - which is not warranted by anything Heim has said, and which she has, moreover, explicitly repudiated; or alternatively, he must see the representational and procedural aspects of DRT as decorative trappings which contribute nothing of substance.

I feel that neither Heim's intentions nor my own are served by such identifications. This is not to deny, however, that the empirical equivalence of the two theories carries important implications. In particular, it carries an important implication for the relation between the particular theory of sections 5-9 and the more general framework which, in section 1, I located at the opposite end of the spectrum of things to which the label DRT has been applied. For, as Heim's

work shows, for the particular predictions in which her theory and the one sketched here converge reference to representations and rules of representation constructions are not needed. This goes to show that we should not be too hasty in taking empirical success as a sign that all the conceptual ingredients of DRT of which I have spoken are in fact indispensible.

If we want further support for these conceptual ingredients, we have to look elsewhere. It is partly with this aim in mind that I now return to the differences between PS and Imparfait which has been our point of departure.

12. TEMPORAL REFERENCE

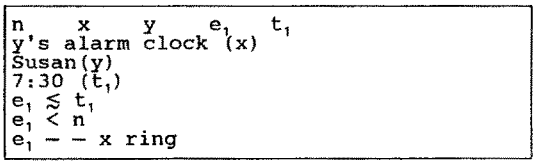
The main assumption underlying the treatment of temporal reference in DRT is that clauses with non-stative verb phrases must be analyzed as descriptions of events, and clauses with stative verb phrases as descriptions of states. The first half of this assumption is a version of the now familiar proposal Davidson made in the sixties concerning the logical form of what he called *action sentences* (Davidson 1984). Within the present framework this means that such clauses introduce drefs that represent events or states. (For mnemonic convenience we will use subscripted "e" 's for event representing drefs and subscripted "s" 's for drefs representing states). With the event and state drefs come several new types of DRS conditions. First, we need conditions which express what type of event or state is represented by a given dref. In current DRT such conditions consist of the dref in question, followed by a DRS, as for instance in:

$$e_1\text{-- } x \text{ ring}$$

Secondly, both new types of drefs enter into conditions which express certain temporal relations. However, in the processing algorithms for tenses and temporal adverbials currently used, these conditions typically involve not only drefs for states and/or events but also drefs representing times (for which we will always use subscripted "t" 's). Among the temporal relations which these conditions must be able to convey are in particular those of precedence, written as $<$; of temporal inclusion, written as $\lesssim$ and of simultaneity, written as $\sim$. Finally, we need a special "indexical" dref, n, to represent the time of utterance.

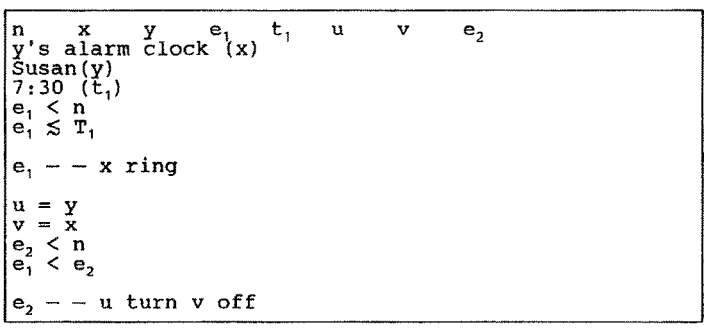
With the help of these new tools, the first sentence of the passage (2) of section III can be represented by a DRS of the following form:[3]

19.



When a sentence like that represented by (19) is followed by another sentence with a non-stative verb phrase in the simple past, this new sentence must often be understood as talking about an event following the one mentioned by its predecessor. This seems to be the case in particular with the second sentence of (2), and an intuitively plausible DRS for the first two sentences of (2), built according to the same principles as (19), looks like this:

20.



As this example indicates, the interpretation of a succession of tensed sentences must often relate the tenses of successive sentences to each other. This is a phenomenon that is somewhat reminiscent of the cases of pronominal anaphora across sentence boundaries that we discussed in sections 7 and 8. However, there is at least one important difference between the two cases. The anaphoric pronouns we have encountered were always interpreted as coreferential with their antecedents. In contrast, in (2) neither the events mentioned in the first two sentences nor

---

[3] About the condition "y's alarm clock (x)", which is to represent the information that x stands for Susan's alarm clock, more ought to be said. There are a number of complications here, partly connected with definite descriptions and partly with the possessive construction Susan's, which I ignore for reasons of presentation.

the occurrence times of these events are understood as being the same. The relation between them is rather that of earlier and later.

For the processing rule associated with the English simple past of non-stative VPs (and similarly for that associated with the PS in French) all this has the following implications. First, as illustrated in (19), the rule must provide for the introduction of a new event-representing dref. Second, it must, under certain conditions,[4] provide for a condition which states that the event represented by the new dref follows upon the event mentioned by the preceding sentence. In particular, in the case of (20) the rule must provide for the introduction of the condition "$e_1 < e_2$".

The interpretation of the remainder of (2) proceeds in essentially the same way, producing a series of drefs representing a temporal succession of events. As the details will not teach us anything that we haven't already encountered in (20) and that is relevant to our present concerns, I will not pursue the DRS construction for (2) any further.

Let us have a look at (3) instead. (3) differs from (2) in that it has not only clauses in the simple past but also a couple that are in the past progressive. It is evident that the temporal relations between that of which these clauses speak and the events mentioned in the discourse surrounding them are different from what we found for the simple past clauses of (2). In particular, the second sentence of (3) describes Susan as being involved in a process that is going on *while* - and not after- the ringing of the phone. This is a general property of the past progressive in English: It is used to convey that a certain process is going on (or, as I prefer to put it for reasons that I

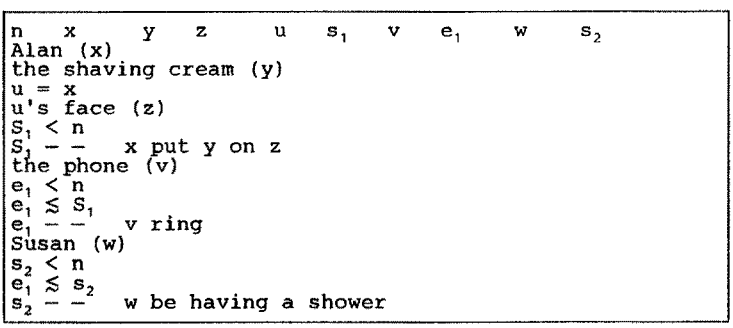
---

[4] The temporal relation that we understand to hold between the events mentioned by two successive sentences in the simple past is not always that of earlier to later. In fact, the temporal relation depends heavily on the "rhethorical" relation between the two sentences. For instance, if the second sentence has the function of explaining what is said by the first the temporal relation between the events is typically the reverse. This is one of the many places where the kind of processing principle on which we are focussing here interacts with aspects of discourse which DRT has only recently begun to explore in earnest. Evidently, rhethorical relations between successive sentences and/or larger discourse units, as well as many other aspects of discourse structure and organization will have to be incorporated into a processing algorithm that is to be able to cope with discourse in a satisfactory way.

cannot go into here, that a certain state obtains) at the time when a certain event, often mentioned in the preceding sentence, occurs.[5]

These observations indicate that the processing rule for the past progressive must differ in at least two respects from that which applies to non-stative VPs in the simple past. First, the new dref which the rule causes to be introduced must be one that represents a state, rather than an event; and second, the rule must provide for a different temporal relation between the represented state and the events which the surrounding discourse mentions.[6] (A rule with similar properties is required for the French Imparfait.) As with the rule for non-stative simple pasts I will leave further details aside and only give, as an illustration of its effects, the DRS for the first two sentences of (3):

21.



---

[5] The past progressive of the first sentence of (3) exemplifies the same principle in a slightly different way; here the state described, i.e. that of Alan putting shaving cream on his face, is understood to have already happened at the time of the event mentioned in the when-clause.

[6] In footnote 4 I noted that the principle of temporal succession for simple past sentences, which is illustrated in (2), is operative only under certain conditions. This might be taken as a concession that the contrast between the simple pasts and the past progressives of non-stative clauses is not as sweeping as the text suggests. It should however be observed in this connection that past progressive and simple past often differ from each other in the way suggested in contexts which are otherwise identical. Compare for instance

"He looked at her. She smiled."

and

"He looked at her. She was smiling."

The difference between past progressive and simple past illustrated in (2) and (3) is of course nothing other than the contrast between the French PS and Imparfait that was summarized in our principle (C2). So construction rules for PS and Imparfait analogous to those we have sketched for the simple pasts and past progressives of English non-statives will account for this contrast (although it may be felt that they do little towards explaining it.) But how could these rules account for the apparently so very different contrast (C1)? The answer is this: In a discourse where the rules for PS and Imparfait work in the way which (2) and (3) illustrate for the English simple past and past progressive, the mentioned events are presented as following each other and, in some instances, as temporally included in certain states. No event emerges as itself including some other event, time, or state. (Formally this will manifest itself in that the event drefs never occur on the right hand side of $\lesssim$.) What this means is that the narrative time emerging from the interpretative process is constituted by a succession of events which, from the perspective of narrative time itself, are all minimal, undivided temporal units.

In contrast, the states introduced by Imparfait sentences will typically have to be interpreted as temporally including certain other temporal elements, e.g. as including events conveyed by clauses in the PS. Take for instance the state represented by s_2 in (21). This state is understood as being already in progress before the time of the event introduced by the preceding sentence. In other words, the state must overlap this event itself as well as some part of the time leading up to it. (In (21) we have omitted to make overlap with the time preceding the event explicit, but strictly speaking the rules for past progressive and Imparfait provide for conditions specifying this; thus the progressive rule yields, in the case of (21), an additional time dref t_1 together with conditions stating that t_1 precedes e_1 and that s_2 temporally includes t_1.)

It must now be clear how the processing rules for PS and Imparfait bear on the principle (C1): If the "punctuality" that (C1) refers to is identified with the property of being undivided in the structure whose constituents are those temporal elements which must be introduced in the process of interpreting a narrative involving PS and Imparfait, then the processing rules associated with these tenses account for why the PS so often conveys punctuality while the Imparfait does not.

While explaining how it is that PS and Imparfait often behave as (C1) would have it, the account they offer should make us wary. We already noted that the processing rules for the tenses do not always work in the way in which they work in the examples we have looked at. So, could it not happen, for instance, that an event which is being introduced by a sentence in the PS gets nevertheless subdivided, say, by some other event that is introduced subsequently?

A little reflection on the ways in which the PS is used shows that this is indeed so. This, however, does not speak against the explanation of (C1) I suggested. On the contrary. In those cases where subsequent discourse contributions entail the subdivision of a previously introduced event we do not perceive that event as punctual, irrespective of whether the event was introduced by a sentence in the PS or in some other way. Thus (C1) is not a principle of absolute validity, but only one that expresses a certain tendency. And the explanation sketched is right not only in that it accounts for the cases which do conform to (C1), but also in that it enables us to see how, under certain conditions, the principle breaks down.

13. CONCLUSION

The DRSs we have just seen involve not only "ordinary" individuals, such as people or sauce pans, but also the intuitively very different categories of events, states, and times. In view of this it might be asked if, in positing such representations, DRT hasn't crossed the boundary that separates first order logic from more powerful logical systems. This is an issue about which there appears to be a certain amount of confusion. (It is a confusion that has nothing to do with DRT, but concerns the logical implications of devices for referring to or quantifying over such entities as events and states irrespective of the precise ways in which these devices are represented.)

In a sense the answer seems perfectly straightforward: The new DRSs do not go beyond first order logic any more than those considered in sections 6-10. In fact, if we write out the truth conditions of the DRSs of the preceding section in the same way in which I described earlier of how one might write out those of the DRSs of sections 6-10, we obtain, once again, expressions that have the unmistakable appearance of first order formulas.[7] (We may recall in this context that Davidson, of whose proposal the treatment of tense sketched here is an elaborate variant, put forward his analysis of action sentences precisely *because* he wanted an analysis of those sentences in terms of the standard predicate calculus (Davidson 1984).)

---

[7] There is a slight complication here, relating to conditions such as e_1 -- v ring. However, these can be translated as atomic formulas of predicate logic if we translate the representation of the relevant verbs as predicates which have an extra argument slot for an event or state.

But haven't we been too hasty? After all, if this was right, then second order logic, in which one quantifies not only over individuals but also over sets of them, would be a kind of first order logic, too. (The only difference between second order formulas and the results of writing out the truth conditions for the DRSs of the preceding section would be that the former have quantification over sets where the latter have quantification over events or states or times.) Yet, the difference between first and second order logic is arguably the most important and profound difference betwen logical systems there is. The difference, of course, is not one of syntactic form, but of *validity*. The set of valid arguments of first order logic is axiomatizable, and we have many different methods of proof, some of which now have remarkably efficient machine implementations, which capture all and only the arguments belonging to that set. But for second order logic validity is provably intractable and no axiomatization is possible.

This consideration should make us pause. Would it be wrong to think of the DRS language of the last section, in whose representations time and temporal relations seem to hold a special position, as determining a notion of validity in which time is an essential ingredient, just as set membership is an essential ingredient in the validity concept for second order logic? There is no conclusive way of deciding this question, I think. But within the present context there is no need to decide it. What matters for us here is that in ordinary reasoning time appears to play an enormously important part. Therefore, if we want to account for a cognitive capacity in which such reasoning does play a part, we better find out, and state, what that sort of reasoning is like. A formalism that will help us with this task, one whose representations are so structured that they enable us to capture the corresponding concept of validity in a correct and implementable fashion, may, for that reason alone, be of considerable value. And it may hold an important advantage over a notational variant which, instead of revealing logically relevant structure, does more to obscure it.

It is for this reason that a system whose representations are translatable into formulas of predicate logic may nevertheless, as a system for knowledge representation, be preferable to predicate logic in its standard form.

To return to the cases we were discussing, a DRS-based system for the repesentation of tensed discourse would certainly commend itself if it provided the basis for a perspicuous and efficient characterization of the logic of time; and it would commend itself irrespective of whether there exists a way of translating its representations into first order formulas.

In the context of time, moreover, such a system, which is capable of proving the inferences belonging to the special logic, would be especially valuable. For, after all, much of our ordinary

reasoning involves inferences that specifically relate to time. So a representation system for temporal discourse that comes equipped with a formal account of what such reasoning involves would catch much that is fundamental to common sense reasoning but that classical logic leaves out.

Indeed, those who, in the course of the past few years, have explored the applications of DRT to the temporal domain are only too well aware of the need for theorem provers that can deal with the logic of time. In fact, such theorem provers are not only needed when DRSs are used as inputs to some inferencing mechanism - as they would for instance in a natural language driven question answering system. They are already needed in the construction of the DRSs themselves. Language interpretation involves its own complex arsenal of inference procedures, and an exact, implementable theory of interpretation must incorporate an account both of what these inferences are and of where and how they enter into the interpretive process.

This, I should add, is not something that is restricted to the temporal domain. We encounter it no matter how we expand DRT - whether it be in the direction of tense, plurals, modality, propositional attitudes, or indirect discourse. As a consequence, no matter which of these different directions we explore, we invariably reach a point after which it becomes impossible to further improve the DRS-construction algorithm if we do not also articulate the underlying forms of inference.

It should be said that on this score DRT has so far been more effective in revealing the gaps than in filling them. But I do not think that this is enough to disqualify it. The design and implementation of new theories of inference, which correspond more closely to the inferences which we ourselves perform (often unconsciously) in the course of exercising a variety of cognitive capacities, is a task which all of us face, and which, I believe, nobody as yet quite knows how to tackle.

It is this, it seems to me, which will be our greatest challenge in the years ahead: To capture the special "logics" of the many concepts that play a prominent role in our ways of thinking and speaking, but which, rightly or wrongly, were not included among the "logical constants" when modern logic cristallized into the predicate calculus as we now know it. Only if that challenge can be met, can we hope to account for meaning and interpretation along the lines which I have sketched. And when that challenge will have been met, (supposing that it can be) DRT - as we know it today - will no doubt seem antiquated, primitive, and puny in comparison with what we will then know and understand.

To concede this is not, however, to dismiss the ideas about discourse, representation, and interpretation in which DRT is grounded and which set it aside from its present competitors. In fact, I am convinced that these ideas are so fundamental to the ways in which our language works that no account of meaning which ignored them would stand a chance.

References

D. Davidson, 1984: The Logical Form of Action Sentences. In D. Davidson: Essays on Actions and Events.

G. Frege, 1879: Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle.

I.Heim, 1982: On the Semantics of Definite and Indefinite Noun Phrases. Ph.D. Dissertation, University of Massachusetts, Amherst.

H. Kamp, 1984: A Theory of Truth and Semantic Representation. In Groenendijk, Jansen, Stokhof (eds.): Truth, Interpretation and Information. Foris.

H. Kamp, 1986: SID without time and question, Forthcoming.

H. Kamp, Ch. Rohrer, 1989: Temporal Reference in French, Forthcoming.

S. A. Kripke, 1972: Naming and Necessity. In D. Davidson, G. Harman (eds.): Semantics of Natural Languages. Dordrecht: Reidel.

R. Montague, 1970: English as a Formal Language. In Visenti (ed.): Linguaggi nella società e nella technica. Milan.