# PRESUPPOSITIONS AND PRONOUNS

**Bart Geurts**

CRIS PI

# Presuppositions and Pronouns

# Current Research in the Semantics/Pragmatics Interface

Series Editors:
**K.M. Jaszczolt,** University of Cambridge, UK
**K. Turner,** University of Brighton, UK

## In this series:

| | |
|---|---|
| **CRiSPI 1 - ed. TURNER** | *The Semantics/Pragmatics Interface from Different Points of View* |
| **CRiSPI 2 - JASZCZOLT** | *Discourse, Beliefs and Intentions* |
| **CRiSPI 4 - ed. JASZCZOLT** | *The Pragmatics of Propositional Attitude Reports* |

## Related books

| | |
|---|---|
| **JASZCZOLT** | *Discourse, Beliefs and Intentions* |
| **JASZCZOLT & TURNER** | *Contrastive Semantics and Pragmatics (2-Volume Set)* |
| **KOERNER & ASHER** | *Concise History of the Language Sciences* |
| **LAMARQUE** | *Concise Encyclopedia of Philosophy of Language* |
| **MEY** | *Concise Encyclopedia of Pragmatics* |
| **TURNER** | *The Semantics/Pragmatics Interface from Different Points of View* |

## Related journals

**Journal of Pragmatics**
*Editor:* Jacob Mey

**Language & Communication**
*Editors:* Roy Harris and Talbot J. Taylor

**Language Sciences**
*Editor:* Nigel Love

**Lingua**
*Editors:* John Anderson and Neil Smith

Free specimen copies of journals available on request

 For further information on the CRiSPI series and for details of how to submit a proposal go to:
*www.elsevier.nl/locate/series/crispi*

# PRESUPPOSITIONS AND PRONOUNS

**Bart Geurts**
*University of Nijmegen*

Current Research in the Semantics/Pragmatics Interface (CRiSPI)

Series Editors: K.M. Jaszczolt, University of Cambridge, UK, and
K. Turner, University of Brighton, UK

The aim of this series is to focus upon the relationship between semantic and pragmatic theories for a variety of natural language constructions. The boundary between semantics and pragmatics can be drawn in many various way, the relative benefits of each gave rise to a vivid theoretical dispute in the literature in the last two decades. As a side-effect, this variety has given rise to a certain amount of confusion and lack of purpose in the extant publications on the topic. This series provides a forum where the confusion within existing literature can be removed and the issues raised by different positions can be discussed with a renewed sense of purpose. The editors intend the contributions to this series to take further strides towards clarity and cautious consensus.

*This page intentionally left blank*

*To Julia*

*This page intentionally left blank*

# Contents

# Preface

In his thumbnail sketch of the history of dynamic semantics, Kamp reports as follows:

> [The study of certain linguistic phenomena] has led to what might be termed a 'dynamic' theory of the semantics of natural language, which was first developed, in at least two different forms, in the early eighties. One of these has come to be known as Discourse Representation Theory (DRT). [...] The semantic approach which DRT exemplifies has been developed also in other ways, most notably in the form of Heim's File Change Semantics, which was conceived at roughly the same time as DRT. [...] Since the early eighties a number of further variants of the approach have appeared, most notably the relational approach of Barwise and the development of Dynamic Predicate Logic of Groenendijk en [sic] Stokhof. (Kamp 1990: 33–34)

This is the standard account. It has it that dynamic semantics is an approach to meaning that was developed in the early eighties by Kamp and Heim, and was subsequently carried on by others, including Barwise and Groenendijk and Stokhof. Although historically speaking there may be some truth to this view, I believe it must be rejected for ideological reasons. In my opinion, Heim's work inaugurated a paradigm which is quite different from the DRT approach. So different, in fact, that I propose to reserve the label 'dynamic semantics' for the approach taken by Heim, Barwise, and the Amsterdam school led by Groenendijk and Stokhof. This approach is truly dynamic in a sense in which DRT is not.

DRT is a theory of interpretation in two senses of the word. It is a theory of meaning and it is also a theory of language understanding. DRT is a cognitivist theory, which is based on the insight that a semantic theory must of necessity take into account the mental processes involved in handling language.

It is somewhat harder to explain in one paragraph what dynamic semantics is, partly because it isn't quite clear what it is. What is clear is that its starting point is the pedestrian observation that utterances cause changes. An

utterance changes the situation in which it occurs, for example, and it causes changes in the interlocutors' cognitive states. It is this observation, I take it, which inspired the notion that the meaning of an expression must be defined in terms of the changes it would cause in a given context. In Heim's words: the meaning of an expression is its context change potential.

At a technical level, DRT and dynamic semantics can be spelled out in such a way that they come to look quite similar, which is presumably what has fostered the myth that they *are* similar, and that any differences between the two theories are to be treated as domestic affairs. I maintain that this impression is quite misleading: the differences between DRT and dynamic semantics are so great that they can hardly be overstated.

One of the ways in which DRT and dynamic semantics diverge is that they give rise to different theories of presupposition projection, which I will call the 'binding theory' and the 'satisfaction theory', respectively. The binding theory, which was first outlined by van der Sandt (1987), is a version of DRT whose main tenet is that presuppositions are entities that want to be bound in the same sense in which anaphors want to be bound. When one thinks in DRT terms, this is a view that comes naturally: although it is theoretically possible to implement just about any account of presupposition in a DRT framework, the binding theory is the most obvious choice. The same can be said of the relation between dynamic semantics and the satisfaction theory of presupposition projection, which goes back to Heim (1983). The satisfaction theory requires that we adopt some version of dynamic semantics, and if one must deal with presupposition projection in a dynamic framework, it is almost frivolous to forgo the satisfaction theory.

The relation between the binding theory and the satisfaction theory resembles that between the semantic theories of which they are offshoots. Prima facie, they are so similar that it has been suggested that they are equivalent, but on closer inspection it appears that they are incompatible. Hence, these two theories and their frameworks arrange themselves in a square of opposition:

$$
\begin{array}{ccc}
\textit{binding} & \leftrightarrow & \textit{satisfaction} \\
\textit{theory} & & \textit{theory} \\
\downarrow & & \downarrow \\
\textit{DRT} & \leftrightarrow & \textit{dynamic} \\
 & & \textit{semantics}
\end{array}
$$

Gloss: 'x $\leftrightarrow$ y' means that x and y are contraries; 'x $\rightarrow$ y' means that x presupposes y as its framework.

The principal aim of this book is to present a version of the binding theory, and apply it to a broad range of data — a *very* broad range, in fact. The data that jointly constitute the notorious projection problem for

presuppositions already involve a rich variety of lexical items and grammatical constructions. But as I see it the projection problem is even bigger than it is generally taken to be, for two reasons. To begin with, the binding theory implies that anaphora is a species of presupposition, so it actually aspires to be a unified theory of presupposition and anaphora. Furthermore, I maintain that presuppositions are crucially involved in the interpretation of modals and attitude verbs, and that, for example, what has come to be known as 'modal subordination' is a quintessential presuppositional phenomenon.

I have a secondary aim, as well. Formulated in the neo-capitalist idiom that has become *de rigueur* in recent years: with a view to boosting the Product's selling power, I will try to show that the Other Product is greatly inferior. That is to say, I will attempt to refute the satisfaction theory and its semantic framework. If I attain these objects I will have made a very strong case for DRT as opposed to dynamic semantics.

The seven chapters of this book fall into three pairs and an extra. In the first two chapters I introduce the main issues concerning presupposition and present the binding theory. Then I assume a rather more polemical posture, and set off on a two-chapter attack on the satisfaction theory and dynamic semantics. In the third pair of chapters my habitual constructive spirit prevails once more, when I apply the binding theory to the interplay between presuppositions and various intensional contexts. The seventh and final chapter presents a presuppositional analysis of names.

This book is a thoroughly revised, rearranged, and expanded version of my 1995 doctoral dissertation. It still counts as a version, I suppose, because it preserves almost all the main ideas of its predecessor; I only changed my mind about disjunctions. The dissertation spawned two articles, which appeared in *Linguistics and Philosophy* (Geurts 1996a, 1998b), and which reappear here, truncated and adapted, as Chapters 3 and 5. Chapters 4 and 7 are based on articles that originally appeared in the *Journal of Semantics* and the *Zeitschrift für Sprachwissenschaft*, respectively (Geurts 1997c, 1997b).

Various people have influenced the form and content of this book in various ways. I should like to thank, first and foremost, Hans Kamp and Rob van der Sandt; and also: Nicholas Asher, Colin Brown, David Beaver, Siegfried Kanngießer, Emiel Krahmer, Manfred Krifka, Marc Ronthaler, Robert van Rooy, and Ede Zimmermann.

*This page intentionally left blank*

CHAPTER 1

# Presupposition

The phenomenon called 'presupposition' gives rise to three main problems, two of which have received much attention in the literature; the third one has been addressed only sporadically. The problem of presupposition failure is the oldest one. If a statement presupposes something which isn't true, what are the consequences for the statement? Will it be false or meaningless, or perhaps wasn't it a statement in the first place? These questions have been debated with so much wit and zest that I am almost reluctant to say that I find them of little interest. This book is not about presupposition failure, it is about presupposition projection, which is the second main problem. When viewed as inferences triggered by certain expressions presuppositions have the remarkable property that, save under rather special circumstances, they tend to go through even when the expressions that trigger them occur in non-entailed positions. This is the projection problem for presuppositions, which is the central theme of this book. The third problem is: Where do presuppositions come from? The standard answer to this question is that presuppositions are triggered by certain words and grammatical constructions. But I will argue that this answer, though perhaps not false, is not good enough.

   The main concern of this book is with the projection problem, and the main purpose of this chapter is to explain that problem, and sketch the outlines of a solution. The projection problem is introduced in § 1.1, where I also propose a diagnostic procedure for determining whether or not something is a presupposition. In § 1.2, I present the intuitive notion of presupposition which underlies the theory that I will present in later chapters. The purpose of the next two sections (§§ 1.3 and 1.4) is to gain a better intuitive understanding of the notion of presupposition by comparing it to two notions that may seem to be intimately related but on closer investigation turn out not to be, viz. conversational implicature and scope. In the last two sections of this chapter I (very) briefly discuss the other two problems mentioned above.

## 1.1  Presupposition projection

Presuppositions are pieces of information which are associated with certain lexical items or syntactic constructions. There are many such items and constructions, and the following is just a small selection:[1]

(1)     *Factives*

    a.    Fred regrets that he cheated at the exam.

    b.    Fred cheated at the exam.

(2)     *Aspectual verbs*

    a.    Barney has stopped writing sonnets.

    b.    Barney has been writing sonnets.

(3)     *It-clefts*

    a.    It was in July that we left for France.

    b.    We left for France.

(4)     *Wh-clefts*

    a.    What Barney ate was potato chips.

    b.    Barney ate something.

(5)     *Quantifiers*

    a.    The Queen has talked with all delegates.

    b.    There were delegates.

(6)     *Definites*

    a.    The pizzeria in the Vatican is closed.

    b.    There is a pizzeria in the Vatican.

Someone who utters any of the (a) sentences commits himself to the truth of the corresponding (b) sentence. Of course, this does nothing to distinguish presuppositions from ordinary entailments. An utterance of (5a), for example, commits the speaker to the assumption that the Queen has talked with all female delegates, but this is an entailment, not a presupposition. The difference becomes apparent, however, as soon as we embed presuppositional expressions or constructions into non-entailing contexts, as in:

(7)     a.    Barney hasn't stopped writing sonnets.

    b.    Maybe Barney has stopped writing sonnets.

---

[1]  The largest collection of presuppositional expressions that I know of is in an unpublished manuscript by Karttunen, from which Levinson (1983: 181–184) quotes extensively.

    c.   If Barney has stopped writing sonnets, his wife will be relieved.

Here (2a) is embedded in the scope of a negation operator, a modal operator, and in the antecedent of a conditional, and, remarkably, it appears that the sentences in (7) commit a speaker to the truth of (2b) just as much as (2a) does. This behaviour sets apart presuppositions from entailments, as the following illustrate:

(8)    a.   The Queen hasn't talked with all delegates.

        b.   Maybe the Queen has talked with all delegates.

        c.   If the Queen has talked with all delegates, she must have a sore throat.

Here (5a) is embedded in the same environments as in the previous example, and as in (7), the presupposition that (5b) is true is part of our understanding of the sentences in (8) just as it is with (5a). In contrast, the inference that the Queen talked with all female delegates, which is an entailment of (5a), is not licensed by any of (8a-c).

    Generally speaking, presuppositions tend to escape from *any* embedded position in the sense that, whenever a sentence φ contains a presupposition-inducing expression χ, an utterance of φ will imply that χ is true. This is only generally speaking, because this rule, though correct in the majority of cases, does not hold without exceptions:

(9)    a.   It is possible that Fred thought the matter over, and that he regrets that he cheated at the exam.

        b.   It is possible that Fred cheated at the exam, and that he regrets that he cheated at the exam.

(10)    a.   If Fred had canard à l'orange, what Barney ate was potato chips.

        b.   If Barney ate anything, what he ate was potato chips.

(11)    a.   Wilma believes that her favorite restaurant in the Piazza di Spagna has gone out of business, and she believes that the pizzeria in the Vatican is closed.

        b.   Wilma believes that there is a pizzeria in the Vatican, and she believes that the pizzeria in the Vatican is closed.

A factive verb like *regret* or *know* triggers the presupposition that its complement clause is true, and so (9a) implies that Fred cheated at the exam (cf. (1a)), which is in conformity with the observation that presuppositions typically escape from embedded positions. However, (9b) shows that this

rule does not always hold. Structurally, (9b) is analogous to (9a), but whereas the latter presupposes that Fred cheated at the exam, the former does not. Similarly with the sentence pairs in (10) and (11). In (10) the *wh*-cleft triggers the presupposition that Barney ate something (cf. (4a)), but unlike the structurally identical construction in (10a), (10b) does not seem to preserve this presupposition. Finally, by uttering (11a) a speaker would normally commit himself to the assumption that there is a pizzeria in the Vatican (cf. (6a)), but an utterance of (11b) would not entail this commitment.

It appears from these observations that presuppositions are normally though not invariably inherited by the sentences in which they occur. This is the so-called 'projection problem' for presuppositions:

> To ask about projection is to ask about the conditions under which presuppositions, no matter how initially triggered, are projected from the clauses in which they are initially introduced to higher-level sentences in which these clauses are embedded. (Chierchia and McConnell-Ginet 1990: 288)

This characterization of the problem, which is representative of what one tends to find in the literature, should be used with caution, because it can be misleading and, moreover, is strictly speaking incorrect, because presuppositions are not carried by sentences. The subject of the verb *presuppose* is a speaker or, perhaps, an utterance. As long as proper care is exerted, it may be convenient to speak of presuppositional expressions as inducing inferences which often do but sometimes do not 'project onto' the sentences in which they occur. But it should be kept in mind that this is just a manner of speaking: whenever it is said that sentence $\varphi$ presupposes that $\chi$, what is actually meant is that, normally speaking, a speaker who uttered $\varphi$ would thereby commit himself to the presupposition that $\chi$ is true.[2]

Compared to other issues in semantics and pragmatics, the projection problem is rather well defined. Speakers' judgments about, say, scope, specificity, donkey sentences, focus, or, for that matter, presupposition failure may be insecure or even inconsistent with one another, but this does not

---

[2]  I am not going to *argue* that presupposition is a pragmatic phenomenon; I take it that this will become evident. The notion of semantic presupposition was a hopeless one from the beginning, and it would probably have withered away somewhere in the seventies if its critics hadn't kept it alive (or, better perhaps, in a comatose state). See (if you must) Kempson (1975), Wilson (1975), Boër and Lycan (1976), Gazdar (1979), Levinson (1983), van der Sandt (1988), Lycan (1984, 1987), and Turner (1992). It is hard to come away from this literature without the lingering suspicion that, considering that it has received so much publicity, there must be something to the notion of semantic presupposition after all. A recent illustration of the trauma this prolonged one-sided debate has caused is Lamarque's 1997 encyclopedia of the philosophy of language, which contains *two* articles on the subject of presupposition, one entitled 'Presupposition, pragmatic', while the other is called simply 'Presupposition'. As if there is a difference.

apply to their intuitions about presupposition projection. I mention this because it might not seem very likely that such a vast and varied array of expressions and constructions should support anything but trivial generalizations. But, likely or not, this expectation proves to be incorrect. There is a non-trivial respect in which aspectual and factive verbs, *it*-clefts and *wh*-clefts, definites, quantifiying expressions, and so on, are alike: they all license inferences which exhibit the projection behaviour that is characteristic of presuppositions. Presupposition is an ubiquitous phenomenon by any standard, and in this book I will argue that it is even more common than has been realized in the past, for instance because anaphora is a species of presupposition.

Despite the fact that the projection problem is reasonably well defined and presuppositions occur all over the place, the history of the notion of presupposition is a troubled one. The beginning of the seventies, in particular, witnessed such fierce debates over this subject that even today some authors will never use the p-word without enclosing it in scare quotes, if indeed they are prepared to use it at all. However, I don't see there is any need to be afraid of presuppositions, and by way of therapeutic service to the less light-hearted, let us consider some of the reasons why presuppositions might be felt to be such dubious creatures. To begin with, there is the sheer variety of, and disparity amongst, theories of presupposition that have been proposed in the past thirty years or so. There are profound differences between, say, the theories presented by Stalnaker (1973, 1974) and Gazdar (1979), and these differences are not merely of a technical nature; the intuitive notions of presupposition that inform these two proposals are at odds with each other. An especially stark illustration of the disparity of the field, at least in its early days, is the work of a Karttunen, who within the span of six years published three theories that were mutually inconsistent, technically as well as conceptually (i.e. Karttunen 1973a, 1974, and Karttunen and Peters 1979).

This extreme dissension is a thing of the past however. In more recent times there has been a growing consensus about what presupposition is. It is widely, though not universally, agreed that presuppositional expressions serve to indicate that some piece of information is taken to be given, and that presuppositions should be accounted for within a so-called dynamic framework, such as discourse representation theory. The two theories of presupposition that take centre stage in this book share this outlook, and between them they account for a major share of the recent publications on the subject. This is not to say that they are equivalent, as has sometimes been suggested. There are differences between the binding theory and the satisfaction theory, and I will argue that they are significant. But they are paltry when compared to the differences between practically any pair of theories proposed in the seventies.

Another reason why the notion of presupposition has a mixed reputation is that it has come in for a considerable share of abuse. It may be hard to find any class of linguistic or para-linguistic inferences that have *not* been dubbed 'presuppositional' at some time or other. This is unfortunate as well as unnecessary, because there is a perfectly adequate procedure for testing if something is a presupposition. This procedure follows directly from the fact that, by definitition, presuppositions are inferences that exhibit projection behaviour. Let us call this procedure the 'Projection Test Battery', or PTB for short.

The PTB consists of three stages. Let $\varphi\{\chi\}$ be a sentence containing a candidate presupposition trigger, which induces the inference that $\chi$ is true. So if $\varphi$ contains the definite NP *the Queen*, then $\chi$ is that there is a Queen; or if $\varphi$ is the *it*-cleft in (12), then $\chi$ is that someone called the police.

(12)      It is Fred who called the police.

In order to establish if $\chi$ is a presupposition, we enter stage one of the PTB: we must check if sentences like the following would normally imply that $\chi$ is true:

(13)      not $\varphi\{\chi\}$
          it is possible that $\varphi\{\chi\}$
          *a* believes that $\varphi\{\chi\}$
          if $\varphi\{\chi\}$ then $\psi$
          either $\varphi\{\chi\}$ or $\psi$

It may be that in some of these environments the inclination to infer that $\chi$ is true (according to the speaker) is stronger than in others; it is sometimes fairly easy to suppress presuppositions triggered in attitude contexts, for example. But at least in the context of a negation or modal operator it should be natural to infer that $\chi$ is true.

(14)      a.      Fred doesn't regret that he kissed Betty.

          b.      It's possible that Fred regrets that he kissed Betty.

          c.      If Fred regrets that he kissed Betty, then she does so, too.

(15)      a.      It isn't Betty who kissed Fred.

          b.      It's possible that it's Betty who kissed Fred.

          c.      If it's Betty who kissed Fred, then he will be disappointed.

It is usually assumed that a factive verb like *regret* triggers the presupposition that its complement is true, and (14) illustrates that this inference passes stage one of the PTB, for it is natural to construe these sentences as implying that Fred kissed Betty. Similarly, it is natural to construe each of the

sentences in (15) as implying that someone kissed Fred, so this inference, which is induced by the *it*-cleft, passes the first stage of the PTB, too.

If our candidate presupposition passes this stage, we come to stage two of the PTB: sentences like the following should *not* imply that $\chi$ is true:

(16)    if $\chi$ then $\varphi\{\chi\}$

it is possible that $\chi$ and $\varphi\{\chi\}$

either not $\chi$ or $\varphi\{\chi\}$

If we apply these tests to factives and *it*-clefts, we get sentences like the following:

(17)    a.    If Fred kissed Betty, then he regrets that he kissed her.

b.    It's possible that Fred kissed Betty, and that he regrets that he kissed her.

(18)    a.    If someone kissed Fred, then it's Betty who kissed him.

b.    It's possible that someone kissed Fred, and that it's Betty who kissed him.

(17a, b) obviously do not imply that Fred kissed Betty, and (18a, b) do not imply that someone kissed Fred, so the presuppositions that are traditionally attributed to factives and *it*-clefts survive the second stage of the PTB, too.

If $\chi$ passes these first two stages, we have established that it normally tends to escape from embedded positions except under certain special circumstances, which is how we defined the notion of presupposition. But if we really want to make sure that $\chi$ is a presupposition we can subject it to a third test. In general, it should be possible to construct special, and therefore marked, instances of the schemata in (13) which block the inference that $\chi$ is true. For example:

(19)    a.    Fred didn't kiss Betty, and therefore he doesn't regret that he kissed her, either.

b.    It's possible that Fred regrets that he kissed Betty, but it's also possible that he didn't kiss her in the first place.

(20)    a.    It isn't Betty who kissed Fred — in fact, Fred wasn't kissed at all.

b.    It's possible that it's Betty who kissed Fred, but it's also possible that he wasn't kissed at all.

Neither (19a) nor (19b) implies that Fred kissed Betty, and neither (20a) nor (20b) implies that Fred was kissed by someone. Hence, factives and *it*-clefts pass the PTB with flying colours. The reader may want to verify that the

same holds for the other presupposition triggers listed at the beginning of this section.

If χ passes all three stages of the PTB, we have strong evidence that it is of a presuppositional nature, and if it fails each single stage we can be sure that it is not. But it should be clear that the PTB is not an effective test procedure; it must not be applied blindly. If a candidate presupposition partly fails the PTB, it may nonetheless be possible to argue that it is a genuine presupposition. For example, van der Sandt (1988: 37-39) points out that negative polarity items will *eo ipso* fail the negation test, and therefore focus particles like *too* would automatically fail the PTB right at stage one.

(21)    a.    Wilma cried, too.

        b.    ?Wilma didn't cry, too.

With focus on *Wilma*, (21a) implies that someone else than Wilma cried. It is widely held that this inference is a presupposition, but the negation test does not justify this assumption, because (21b) is simply infelicitous. However, we shouldn't conclude from this that *too* is not a presupposition trigger, because its alleged presupposition passes all stages of the PTB *provided* we leave negative contexts out of account; and there are good grounds for relaxing the PTB in this case, precisely because *too* is a negative polarity item. This example shows quite clearly that, when testing for presuppositionhood, we shouldn't rely on the negation test alone, as many people have done, and some still do. In fact, we shouldn't rely on any single test, if only because there is no reason to: the term 'presupposition projection' stands for a rich tapestry of regularities, and our diagnostic procedure should reflect this. Thence the PTB.

Rich though the notion of presupposition may be, it should be acknowledged that it has sometimes been pushed too far. But if such a large number of phenomena have been dubbed 'presuppositional', it is mainly due to the fact that the available diagnostics are rarely applied with sufficient rigour. Let me give an example. Usually, a *wh*-question like (22a) will license the inference that the speaker takes (22b) to be true:

(22)    a.    Who killed Caesar?

        b.    Someone killed Caesar.

It has often been claimed that this inference is of a presuppositional nature, i.e. that (22a) presupposes (22b).[3] However, this claim is hard to assess, because by and large the PTB doesn't apply to questions. For example, the standard negation test breaks down when applied to questions. It is true that

---

[3]    See, for example, Keenan and Hull (1973), Hintikka (1983), and, for a critical discussion of this position, Groenendijk and Stokhof (1997).

there is a way of negating (22a) that does not preserve the inference that the speaker takes (22b) to be true:

(23)    Who didn't kill Caesar?

But this does not prove that (22a) doesn't presuppose (22b). If (22a) triggers the presupposition that (22b) is true, it is presumably the entire question that acts as a presupposition inducer. So when applying the negation test, we have to determine, not what happens when we insert a negation somewhere *into* this alleged presupposition inducer but rather what happens when a negation sign takes scope over the question as a whole. But of course, questions are allergic to external negation:

(24)    ?It is not the case that who killed Caesar?

If we want to prove that (22a) presupposes (22b), then it must be shown that this inference exhibits projection behaviour, and that turns out to be difficult, because as a rule questions cannot occur in embedded positions. There may be one or two exceptions though.

(25)    a.    Fred asked who killed Caesar.

        b.    If Brutus wasn't in town, then who killed Caesar?

        c.    If Ceasar was killed, then who killed Caesar?

In (25a), an interrogative sentence occurs in an embedded position, and if it were the case that we would normally infer from this sentence that someone killed Caesar (which isn't too clear), then we would at least have some evidence that this is a presupposition triggered by the *wh*-complement. The contrast between (25b) and (25c) is more significant, because in the former case there is a strong tendency to infer that, according to the speaker, someone killed Caesar, which is absent in the latter case. This is precisely what we should expect if (22c) triggered the presupposition that (22b) is true. However, this is the only halfway sound piece of positive evidence that we have turned up so far, and even it is not beyond suspicion, because the contrast between (25b, c) only counts if the consequent of these conditionals is equivalent in all relevant respects (whatever they may be) to the sentence in (22a). This might entail that (25b, c) are conditional questions, and there has been quite some resistance against the notion of conditional speech act, which is pointedly summarized in Walker's (1975: 145) rhetorical question: 'How can one perform a speech act in a conditional way, any more than one can stand on one's head in a conditional way?'

To sum up: there is precious little direct evidence for the hypothesis that a *wh*-question presupposes the corresponding existential declarative, because it is so difficult for questions to occur in embedded positions, if they can occur in such positions at all. In other words, it is a non-trivial matter to show

that questions give rise to inferences that exhibit projection behaviour. However, the chief purpose of the foregoing discussion was not to establish that this is so. Its purpose was mainly to indicate the issues that need to be addressed if one wants to test the hypothesis that questions are presupposition inducers, and I believe this was a useful exercise because, to the best of my knowledge, *none* of the authors who have considered this hypothesis have addressed these issues in any systematic way.

If the PTB had always been wielded with proper care, many proposals for granting this or that expression presuppositional status would not have been published, but nevertheless there remains a large and in many respects heterogeneous class of expressions and constructions of which there can be no doubt that they are presupposition inducers. It is sometimes suggested that the mere fact that this class is so large and so diverse makes it unlikely that the notion of presupposition is a useful one. If they are so many and so different, whatever it is they have in common is bound to be something utterly trivial.[4] This suggestion is misplaced. Consider the verbs of English. If the class of presupposition triggers is large, then the class of English verbs is huge, and surely it isn't less heterogeneous, either. Therefore, it is a priori unlikely that the class of English verbs will yield anything but trivial generalizations. This is evidently a non sequitur. But from a logical point of view, it will not become less invalid if 'verb' is replaced with 'presupposition trigger'.

There is one further reason why the concept of presupposition may be felt to be suspect. It is that presuppositional phenomena bear some resemblance to certain other semantic and/or pragmatic phenomena. On the one hand, it has been claimed that presuppositions are related to, and perhaps even a species of, conversational implicatures. On the other hand, what may be informally described as the tendency of presuppositions to take wide scope has led some people to maintain that presupposition projection can be treated in terms of scope. I turn down both hands, and in §§ 1.3 and 1.4 I explain why.

## 1.2  Taking something for granted

Here is what Longman's dictionary of contemporary English has to say about presupposition:

> **pre·sup·po·si·tion** /ˌpriːsʌpəˈzɪʃən/ *n* **1** [U] the action of presupposing (PRESUPPOSE) **2** [C, C5] something that is PRESUPPOSEd: *Your*

---

[4]  For statements to this effect, see Boër and Lycan (1976), Levinson (1983: 217), Lycan (1987), and Neale (1990: 54), among many others.

*judgment of the facts is based upon the presupposition that they are*
*true; such presuppositions are unwise.*

This may not seem to be very instructive, yet it is useful to be reminded that
*presupposition* is a deverbal noun, and that one speaks of presuppositions in
order to characterize not sentences or words but people and what they do.
Presupposition, according to Longman's dictionary, is either the action of
presupposing or its object. But what does it mean to presuppose something?

> **pre·sup·pose** /ˌpriːsə'pəʊz/ *v* [T1, 5] **1** to suppose or take to be true
> without trying to find out: *A scientist never presupposes the truth of*
> *an unproved fact* **2** to be necessary as something that comes before
> (something else) according to reason: *A child presupposes a mother.*
> | *An honour given to a person presupposes that he has earned it.*

Now, although I hold Longman's dictionary in high esteem, it seems to me
that the second sense listed here is marginal at best. A child doesn't
presupposes a mother any more than a sentence presupposes a sentence. Be
that as it may, it is evidently the first meaning that is the primary one, and it
says that to presuppose something is to take it for granted. Again,
presupposing is something people do, and it is not necessarily verbal
behaviour, although it may be that it is typically expressed by linguistic
means.

   This folk concept of presupposition is the basis for the theory that I present
in this book. The theory has its roots in a triad of papers by Stalnaker
published in the first half of the seventies (i.e. Stalnaker 1970, 1973, and
1974). In keeping with the layman's view, Stalnaker's starting point is that
presupposition is a 'propositional attitude', which is to be understood in
behaviourial terms:

> A speaker presupposes that P at a given moment in a conversation
> just in case he is disposed to act, in his linguistic behavior, as if he
> takes the truth of P for granted, and as if he assumes that his
> audience recognizes that he is doing so. (Stalnaker 1973: 448)

To presuppose something is like making a promise. If I promise my daughter
to buy her an ice cream, it is completely irrelevant what my actual beliefs and
intentions are: I made a promise regardless whether I intended to keep it or
not. Likewise, a speaker who presupposes something incurs a commitment
(to use Hamblin's expression) regardless whether he really believes what he
presupposes. To be sure, presuppositions are normally believed to be true,
but belief is not a prerequisite for presupposition.

   Presupposition is the dual of assertion, and if the former is a propositional
attitude, then so is the latter, and in the same sense. Like presupposition,
assertion implies commitment, not true belief. If a story teller asserts that the

princess kissed a frog, he is committed to the truth of his statement for the duration of his story; but we don't expect him to believe what he says. The difference between presupposition and assertion is that what is presupposed is taken for granted, and what is asserted is not. More accurately, to presuppose something is to represent oneself as assuming that the presupposition is already part of the common ground of assumptions that the interlocutors share between them.[5] A presupposition is *presented as* 'an item of presumed common knowledge' (Stalnaker 1973: 450), and just as one can present old stories as hot news, a speaker can present new information as if it were already part of the common ground. In such a case, the speaker dispenses new information by pretending that his audience already know. Normally, this pretence will be a transparent one, and it will be recognized by everyone that this is what is going on. 'In some cases, it is just that it would be indiscreet, or insulting, or tedious, or unnecessarily blunt, or rhetorically less effective to openly assert a proposition that one wants to communicate.' (Stalnaker 1974: 202) Borrowing an example from Gazdar (1979: 106), if I arrive late at a meeting, I may say by way of excuse:

(26)    I'm sorry I'm late, my car broke down.

By saying this I presuppose that I came by car. It may be that no one in the audience knew this beforehand, but they will let me get away with my presupposition because it is innocent (it is not unusual for people to drive to meetings), and because they appreciate that in order to avoid presupposing that I came by car I would have had to use a prolix formula such as:

(27)    I'm sorry I'm late, I came by car and it broke down.

Hence, new information can, and often will, be conveyed by way of presupposition, but it is important to realize that this is an instance of what Grice has called 'exploitation': the speaker exploits the rules of conversation by breaking them.

   Lewis (1979) has given this form of exploitation a name which has gained currency in the literature: 'accommodation'. But although Lewis endorses Stalnaker's theory of presupposition, the way he characterizes accommodation is less circumspect than Stalnaker's, and this may have contributed to the considerable variety of misinterpretation and abuse that the concept of accommodation has come in for. Lewis says that:

> [...] it's not as easy as you might think to say something that will be unacceptable for lack of required presuppositions. Say something

---

[5]  Stalnaker uses the term 'presupposition set' to refer to these shared assumptions. I will not adopt his usage in this point, because I think it is potentially confusing. I will restrict the notion of presupposition to what a speaker does by way of making an utterance. This is just a terminological difference though.

> that requires a missing presupposition, and straightway that presupposition springs into existence, making what you said acceptable after all. (Lewis 1979: 339)

This is misleading, to the say the least.[6] Contrary to what Lewis maintains, it is perfectly easy to say something that is unacceptable for lack of required presuppositions. For example, if in the situation discussed above I had said (28) instead of (26), my audience would not have let it stand, for obvious reasons.

(28)     I'm sorry I'm late, my chariot broke down.

If speakers often get away with presupposing something that is not shared knowledge, it is because they have a finely developed sense for what information can or cannot be conveyed in this manner. It is not because accommodation is a rule that is applied by default or even automatically.

Karttunen, too, had the notion of accommodation some years before Lewis gave it its name. Karttunen gives the following examples (1974: 191):

(29)   a.   We *regret* that children cannot accompany their parents to commencement exercises.

       b.   There are *almost* no misprints in this book.

       c.   John lives in *the third brick house down the street from the post office*.

       d.   It has been *pointed out* that there are counterexamples to my theory.

The italicized expressions trigger presuppositions that can be interpreted by way of accommodation. In (29a), for instance, the factive verb *regret* triggers the presupposition that children cannot accompany their parents to commencement exercises, and it is this clear that this sentence may be used felicitously in a context in which the presupposition is not yet part of the common ground. In fact, a construction like this will often be preferred to a straighforward assertion like (30), which may sound too blunt:

(30)     Children cannot accompany their parents to commencement exercises.

Accommodation is often described as a repair strategy. This is not the way Stalnaker puts it, but that is because he prefers to take the speaker's point of

---

[6]   To be fair to Lewis it must be noted that he adds as a parenthesis that 'at least, that is what happens if your conversational partners tacitly acquiesce'. If in the subsequent literature this crucial rider was sometimes forgotten, it is not entirely Lewis's fault, although he is to blame for de-emphasizing it.

view. When we adopt the hearer's perspective, however, it becomes natural to view accommodation in these terms. If the speaker presupposes something that is not yet part of the common ground, then the hearer may be prepared to go along with the speaker's pretence that his presupposition is already given, and revise his representation of the common ground accordingly. Viewing accommodation from the hearer's point of view helps to bring out and motivate the following constraint on interpretation. Given that accommodation is a repair strategy, it is reasonable to suppose that the hearer will attempt to construe the speaker's utterance in such a way that accommodation is required as little as possible. In other words, the hearer will try to deal with the speaker's presuppositions, if at all possible, while keeping his representation of the common ground fixed (and, of course, it is mutual knowledge between the interlocutors that this is what the hearer will do, and the speaker will devise his utterances accordingly). This principle will play an important role in later chapters.

In the previous section we defined presuppositions as inferences that are triggered by certain expressions, and that exhibit projection behaviour. For Stalnaker, a presupposition is an assumption which a speaker takes for granted. These two definitions may seem to contradict each other, but the contradiction is only an apparent one. For we can plausibly say that a given form of words requires that the speaker presuppose something or other. To illustrate, a speaker who chooses to use a factive construction is thereby obliged to presuppose that its sentential complement is true. Or, in Stalnaker's own terms, certain expressions impose 'presuppositional requirements' on the speaker who wants to use them. This view on the relation between presuppositional expressions and what a speaker presupposes is attractive because it allows for the possibility that not all presuppositions are of lexical or grammatical origin. Later on in this chapter I will consider one way of exploiting this leeway.

We now come to what in my opinion is the *pièce de résistance* of Stalnaker's work on presupposition, which is his sketch of how the pragmatic theory of presupposition might solve the projection problem. To explain this, let us focus on the projection behaviour of presuppositions triggered in conjunctions and conditionals.

(31)  a.  It was Fred who shot the porter, and Barney took care of the manager.

  b.  If it was Fred who shot the porter, then Barney took care of the manager.

(32)  a.  Barney took care of the manager, and it was Fred who shot the porter.

     b.    If Barney took care of the manager, then it was Fred who shot the porter.

(33)    a.    Someone shot the porter, and it was Fred who shot the porter.

     b.    If someone shot the porter, then it was Fred who shot the porter.

Each of these sentences contains a cleft sentence which triggers the presupposition that someone shot the porter, but while a speaker who uttered any of the sentences in (31) or (32) would be required to make this presupposition, he would not be taken to presuppose this if he uttered either of the sentences in (33). In the case of (33a) we would say that it had been asserted rather than presupposed that someone shot the porter, and (33b) doesn't imply in any way that the porter has been shot. Intuitively it is obvious why the presupposition is blocked in (33) but not in (31) or (32). Apparently, if the second half of a conjunction or conditional is of the form $\varphi\{\chi\}$, and $\chi$ is entailed by the first half, then the presupposition will be blocked; otherwise it will go through. So the following generalizations seem to hold:[7]

(34)    a.    '$\varphi\{\chi\}$ and $\psi$' requires the presupposition that $\chi$

          '$\varphi$ and $\psi\{\chi\}$' requires the presupposition that $\chi$, unless $\chi$ is entailed by $\varphi$

     b.    'if $\varphi\{\chi\}$ then $\psi$' requires the presupposition that $\chi$

          'if $\varphi$ then $\psi\{\chi\}$' requires the presupposition that $\chi$, unless $\chi$ is entailed by $\varphi$

Suppose now that it is suggested that these are not just empirical generalizations, but lexical rules: (34a) is part of the lexical meaning of *and*, and (34b) is part of the lexical meaning of *if . . . then*. There are several problems with this suggestion. First, if we just postulate that (34a, b) are lexical rules, we don't have much of an explanation to boast of; we have merely codified our observations in the lexicon. Secondly, if we just say these generalizations are accounted for by lexical rules, and leave the matter at that, it is implied that these are purely conventional regularities, and that doesn't seem to right, since it is intuitively obvious that these generalizations hold for a reason. Thirdly, and this is what Stalnaker (1973, 1974) worries about, this proposal entails that the meanings of the connectives are much more complex than they ought to be, intuitively speaking. In particular, it

---

[7] These are Karttunen's (1973a) initial filtering conditions for *and* and *if . . . then*. Karttunen shows that these conditions are not quite adequate, and ends up proposing rather more sophisticated rules, but these fine points need not concern us here.

implies that *and* is non-commutative ('φ and ψ' no longer has the same meaning as 'ψ and φ'), which Stalnaker considers an unpalatable consequence, and rightly so.

An utterance changes the context in which it is made. If I say,

(35)      The struggle of class against class is a political struggle.

the context will change in a number of ways. For one thing, I have committed myself to the claim that this statement is true. For another, if my audience indicate that they accept my statement (by nodding their consent, for example, or by tacit acquiesence), then it becomes part of the common ground that (35) is true. Stalnaker's crucial insight is that such changes may take place midway in an utterance, and that these intermediate context shifts are the key to the solution of the projection problem.[8]

A speaker who utters a sentence of the form 'φ and ψ' in a context c, changes c in two steps (assuming that his audience accept what he says): first c is enlarged to a context c' in which it is part of the common ground that φ is true, and then c' is enlarged to a context in which it is part of the common ground that ψ is true. Hence, the presuppositions required by φ and ψ arise in *different* contexts: presuppositions required by φ are taken to be given in c, while presuppositions required by ψ are taken to be given in c'. Presuppositions triggered in ψ can still impose restictions on c, but only indirectly. Since c' = c-plus-φ, whatever is presupposed in c' and not entailed by φ will be presupposed in c. But any presupposition triggered in ψ that is entailed by φ will impose no restrictions on c. In brief, the presuppositional asymmetry of conjunctive sentences is due to the obvious fact that a speaker who is in the process of uttering such a sentence will take for granted the truth of the first conjunct when he starts uttering the second (unless his interlocutors express their disagreement).

This account can be extended for dealing with conditional sentences as follows. It is plausible to assume that the antecedent of a conditional statement of the form 'if φ then ψ', as uttered in a context c, serves to *temporarily* change c into a context c' in which φ holds; for this is what it means to suppose that φ is true, which is what the *if*-clause does. Therefore, the consequent and its presuppositions occur not in c but in c', and any presuppositions required by ψ that are entailed by φ will not be perceived as presuppositions of the conditional statement as a whole. Conjunctions and

---

[8]   Karttunen's 1974 theory is based on the same insight, but there is an important difference between Karttunen's position and Stalnaker's. Karttunen explicitly rejects the notion of speaker presupposition, and although his is not a semantic theory of presupposition, it does lean rather heavily towards a semantical approach. The satisfaction theory, which evinces the same semanticist bias, has taken on Karttunen's legacy, not Stalnaker's. See Chapter 3 for further discussion.

conditionals have the same projection profile because both types of statement affect their contexts in two consecutive steps.

This is Stalnaker's pragmatic account of presupposition projection. Like the wheel, it is based on a simple but compelling idea, and it is just too good to be wrong. There are two aspects of Stalnaker's theory that I want to emphasize. First, it is a pragmatic theory. Sentences don't change contexts; they don't do anything. It is speakers who employ sentences to change the context. This distinction is an important one because the past two decades have witnessed the rise of a paradigm, inaugurated by Heim's (1982, 1983) work on anaphora and presupposition, which incorporates Stalnaker's insights by casting them in a semantical mould. These developments will be discussed at length in later chapters, but I should like to note beforehand that, although they were inspired by Stalnaker's work, they depart from his intentions in a quite radical way. By contrast, the theory that I advocate remains loyal to Stalnaker's concept of speaker presupposition.

The second point that bears emphasizing is that on Stalnaker's account of presupposition projection, it need not be assumed that presuppositions are cancellable. We have seen that a presupposition triggered in the consequent of a conditional seems to disappear if it is entailed by the antecedent, and it is sometimes said that in such cases presuppositions are eliminated or cancelled. In some theories of projection this is not merely a manner of speaking, since they require that presuppositions are literally cancellable.[9] On Stalnaker's account presuppositions are never cancelled. If they sometimes seem to disappear it is because they may be satisfied by a local context. Thus, if $\varphi$ entails $\chi$, the presupposition triggered in 'if $\varphi$ then $\psi\{\chi\}$' seems to 'disappear' because in the intermediate context in which $\psi\{\chi\}$ is evaluated it is already given that $\chi$ is true. Nowadays it is widely, though not universally, agreed that this view is correct, and that, strictly speaking, presuppositions can never be cancelled.

In the following chapters, I present a theory which is based on Stalnaker's insight that presuppositions are pieces of information which are taken to be given, either in the main context or in some local context. However, what I have to offer is not just a theory of presupposition but a unified account of presupposition and anaphora. That is to say, in the following I will adopt and elaborate upon van der Sandt's (1987, 1992) hypothesis that anaphora is a species of presupposition. This hypothesis will be discussed at some length in the next chapter; the following examples merely are to give it some initial plausibility:

(36)    a.    Fred cheated at the exam, and he regrets {it/that he cheated at the exam}.

---

[9] Theories that require presuppositions to be defeasible have been proposed by Karttunen (1973a), Gazdar (1979), van der Sandt (1988), Mercer (1992), and many others.

> b.    There were delegates from all provinces, and the Queen talked with all {of them/delegates}.
>
> c.    There is a pizzeria in the Vatican, and {it/the pizzeria in the Vatican} is closed.

In each of these examples, a pronoun and a presuppositional expression perform the same duty: if the neuter pronoun in (36a) picks up an antecedent introduced in the first conjunct, it is natural to suppose that the *that*-clause does exactly the same thing. Analogous remarks apply to (36b) and (36c).

These observations bring out an aspect of presupposition that we haven't touched upon so far, because it doesn't figure explicitly in Stalnaker's writings.[10] We have said that if a speaker presupposes that $\chi$, he takes it for granted that $\chi$. This is not quite correct, however, or at any rate it is incomplete. A presupposition is not just something that is taken to be true in the given context; it is something that is retrieved from the context. The distinction is subtle but consequential. It is perhaps easiest observed in definite NPs. We have said, as is common in the presupposition literature, that a definite NP of the form 'the N' triggers the presupposition that there is an N. (36c) shows that this is not all, and that it is better to say that the function of such an NP is to retrieve an N from the current context. Similarly, it is true that the quantified NP *all delegates* in (36b) requires that there be delegates, but this is not enough to do justice to the presuppositional requirements of this NP. Rather, what we should say is that the presuppositional function of this NP is to retrieve from the context some set of delegates. Another case in which the distinction is especially vivid are the presuppositions triggered by focus particles such as *too*:

(37)    I am hungry, too.

With focus on the subject term, the particle *too* triggers the presupposition that there are persons other than the speaker who are hungry. But this much can be taken for granted in just about any context, and it is clear that (37) can only be used in contexts in which it is clear whom the speaker has in mind, for example when someone else has previously uttered the same sentence without the *too*.

The pronoun *it* in (36c) does not merely convey that this statement is about a non-human individual; it is an instruction to the hearer to retrieve (his representation of) the individual in question. Likewise, the definite NP *the*

---

[10] Which is not to say that it is inconsistent with Stalnaker's views. If I were bent on proving that Stalnaker's work prefigures every major aspect of the theory that I am going to defend, I would certainly point to some of his remarks about pronouns, which show that he views anaphors as presupposition-inducing expressions in some sense (cf. Stalnaker 1973: 449, 454). But the sober truth is that nothing in Stalnaker's writings indicates a clear awareness of the intimate relationship between presupposition and anaphora.

*pizzeria in the Vatican* in (36c) does not merely convey that there is pizzeria in the Vatican; it is an instruction to the hearer to retrieve (his representation of) the said restaurant. Anaphoric pronouns, on this view, are just semantically attenuate presuppositional expressions, and the theory of anaphora is just one of the departments of the theory of presupposition. In a nutshell, this is the binding theory of presupposition.

### 1.3 Presuppositions vs. implicatures

When taken at face value, presuppositions appear to be defeasible inferences. By definition, presuppositions are *normally* impervious to any form of embedding, and will be suppressed under special circumstances only. We have seen that this impression is actually misleading, because in reality presuppositions are never cancelled; they may be intercepted in an intermediate context, but they never disappear altogether. However, let us ignore this for a while, and suppose that presuppositions are cancellable, as they seem to be. If this is so, it is something presuppositions have in common with conversational implicatures, and it is not to be wondered that many people have considered the possibility of explaining presuppositions along Gricean lines. One of the first authors to explore this possibility was Grice himself, although, characteristically, Grice does not align himself with the position (as he puts it) that presuppositions are implicatures (Grice 1981). He entertains the hypothesis that some alleged instances of presupposition might be partly explained in terms of implicature, but also considers the possibility that presupposition might be partly explained in terms of scope. In short, Grice doesn't risk taking a stance. Where Grice feared to tread, others have been more daring. For instance, Atlas and Levinson (1981) argue that the alleged presuppositions of *it*-clefts are in fact conversational implicatures, and it has often been suggested that definiteness and/or anaphora might somehow be analysed in terms of implicature (see Levinson 1987, 1997, and Gundel et al. 1993).

In keeping with this trend, it has been claimed that there is such a thing as a projection problem for conversational implicatures. This view is at the heart of Gazdar's (1979) theory of 'implicature projection', and it is enunciated quite explicitly by Atlas and Levinson, too:

> The 'filtering' of implicata seems to occur in compound sentences in the same way that filtering of presupposition occurs. (Atlas and Levinson 1981: 5)

More recently, this claim was reiterated by Levinson (1997). To be sure, it is not necessary for a would-be Gricean reductionist to accept this position, but

the main goal of this section is not to argue against any particular theory, but to delineate presuppositions from implicatures, and one way of doing this is by showing that there is no such thing as a projection problem for implicatures.[11]

Presuppositions are triggered by certain lexical items and grammatical constructions, and therefore it makes sense to say that presuppositions may be induced in embedded positions. Conversational implicatures, in contrast, are licensed by *utterances* in certain contexts, and are not, or at least not necessarily, tied up with specific linguistic forms. Therefore, it is a category mistake to speak of conversational implicatures as arising in embedded positions.[12] To bring home this point, let us compare the following examples:

(38)     a.     Fred kissed some of the girls.

         b.     It is not the case that Fred kissed all the girls.

(39)     a.     Wilma believes that Fred kissed some of the girls.

         b.     It is not the case that Wilma believes that Fred kissed all the girls.

The definite NP *the girls* triggers the presupposition that there were girls, and sure enough both (38a) and (39a) inherit this presupposition, for normally speaking either one of these statements would be taken to imply that a group of girls is contextually given. At the same time, these sentences seem to somehow imply the falsity of (38b) and (39b), respectively.[13] These inferences may be construed as conversational implicatures, along the following lines. It is plausible to suppose that the truth conditions associated with *some* are such that (38a) is strictly speaking true even in a situation where all the girls were kissed by Fred. An utterance of (38a) in such a situation might be misleading but it would not be false. Nonetheless, the

---

[11]  Sadock (1978) and Karttunen and Peters (1979) have proposed that (some) presuppositions are really conventional implicatures, but, as Levinson (1983: 217) observes, since the concept of conventional implicature is such as troubled one, this is just a terminological sidestep, and I will not consider it further. In general, when I speak of 'implicatures' *tout court*, I will mean conversational implicatures.

[12]  It might be argued that only *generalized* conversational implicatures show projection behaviour (cf. Levinson 1997). Since, according to Grice (1975), a generalized conversational implicature *is* associated with 'a certain form of words' (which I take to imply that it is a conventionalized conversational implicature), it would at least be conceptually coherent to claim that implicatures can be triggered in embedded positions. But otherwise this position suffers from the same problems that I will discuss in the following.

[13]  In the case of (39a) it is tempting to infer that Wilma believes that Fred didn't kiss all the girls, which is stronger than (39b). But this inference can be left out of account here.

recipient of this sentence will usually be entitled to conclude that the sender takes (38b) to be false, because he may reason as follows: (i) Presumably, the speaker knows what he is talking about, and therefore he knows how many girls Fred kissed; (ii) if the speaker had known (38b) to be true, he would surely have produced this sentence instead of (38a); but (iii) since he hasn't done that, it follows from (i) and (ii) that he takes (38b) to be false.

If some such story can be told about (38a), a strictly parallel story can be told about (39a), which embeds (38a). Here is how: (i) Presumably, the speaker knows what he is talking about, and therefore he knows how many girls Wilma believes were kissed by Fred; (ii) if the speaker had known (39b) to be true, he would surely have produced this sentence instead of (39a); but (iii) since he hasn't done that, it follows from (i) and (ii) that he takes (39b) to be false. Note that this account is given without reference to the circumstance that in (39a), the quantified NP *some of the girls* occurs in an embedded position. The crucial premise for this line of explanation is that (38a) and (39a) become stronger statements when *all* is substituted for *some*; the notion of embedding doesn't enter the picture in any way.

There is a further reason why it is wrong to describe the inferences in (38) and (39) as instances of implicature projection. If these really were instances of projection in the relevant sense, then the inference from (39a) to (39b) would be beside the point, since the crucial prediction would be that (39a) will normally imply (38b). For, after all, projection means that the *same* inference tends to through in all sorts of linguistic environments. But evidently (39a) does not normally imply that (38b) is true.

There are cases that might be construed as instances of implicature projection, but only when taken at face value. The following is such a case:

(40)    It is possible that Fred kissed some of the girls.

This will tend to be heard as implying that, according to the speaker, Fred didn't kiss all the girls. But, again, this inference can be accounted for, along the lines set out by Grice, without anything like a projection mechanism entering the scene. Clearly, (40) is weaker than:

(41)    It is possible that Fred kissed all of the girls.

So if a speaker utters (40) instead of (41), he presumably takes the latter to be false, which entails that he doesn't believe that Fred kissed all the girls, i.e. he believes that (38b) is true. Again, no need for a notion of implicature projection. Furthermore, the inference vanishes without a trace if we replace the epistemic modal with, say, a deontic one:

(42)    Fred was given permission to kiss some of the girls.

A hearer who takes this to imply that (38b) is true must be assuming that Fred is an obedient boy, but it is clear that this will not be inferred by default.

Note, by contrast, that the presupposition that a group of girls is contextually given, which is triggered by the definite NP *the girls*, will go through no matter what.

For good measure, here are some further examples that demonstrate how presuppositions and implicatures diverge:

(43)     Wilma hopes that
         Wilma doubts that    $\left.\right\}$ Fred kissed some of the girls.
         Can you tell me if
         Let's go and see if

According to my intuitions, none of these sentences definitely imply that Fred didn't kiss all the girls, and at least some of them definitely don't imply this. By contrast, each of these sentences clearly implies, in the absence of evidence to the contrary, that there were girls. Presuppositions show projection behaviour; implicatures do nothing of the sort.

If it cannot be shown that conversational implicatures are liable to escape from embedded positions, as presuppositions are, then *a fortiori* it cannot be shown that they are blocked or suspended under some circumstances. Cases like the following, for example, do not show that some implicatures exhibit projection behaviour:

(44)     a.    The water is warm, and perhaps even hot.

         b.    The water is warm. In fact, it is hot.

         c.    The water is warm, if not hot.

Horn (1989: 234-235) makes a distinction between implicature blockers and implicature suspenders. In (44a) an implicature is blocked, according to Horn: in this case the speaker explicitly conveys that, for all he knows, the water may be hot, and therefore the implicature that the water is not hot, which would normally be licensed by the use of the scalar adjective *warm*, is blocked. In (44b) this implicature is suspended: the first sentence licenses the implicature, but it is removed by the second statement, which contradicts it. Finally, (44c) is ambiguous and may be construed as an instance either of suspension or blocking. In the former case, the speaker is taken to convey that he doesn't know if the water is hot; in the latter, his statement implies that the water is not hot.

Judging from the way he describes what is going on in (44a-c), Horn must be assuming that the quantity implicature induced by *warm* exhibits projection behaviour. But it doesn't really. Consider (44b) first. Here the implicature is licensed by a simple sentence, only to be cancelled by the next statement. This could never happen to a presupposition. Compare (44b) with:

(45)    ?Fred's wife did it. But Fred isn't married.

In general, presuppositions triggered in entailed positions, such as simple sentences, will be entailments; they cannot be cancelled. Conversational implicatures, by contrast, are always cancellable. (44a) is similar in this respect to (44b). According to Horn, the implicature is suspended in this case, and again this is something that cannot be done with presuppositional inferences:

(46)    ?Fred's wife did it, and perhaps Fred isn't married.

Finally, consider (44c). Horn's description suggests that, on one reading of this sentence, an implicature triggered in the consequent of the conditional is suspended in the antecedent, just like a presupposition might be suspended. That it to say, it is suggested that (44c) is analogous to (47a):

(47)    a.    If Fred is married, then Fred's wife did it.
        b.    If today is Tuesday, then Fred's wife did it.

If this analogy is correct, then we should expect that the implicature is not suspended if the antecedent of the conditional doesn't entail it; for this is how presuppositions behave, as witness (47b), which does imply that Fred has a wife. This is not how implicatures behave, though. Cf.

(48)    If today is Tuesday, then the water is warm.

This clearly does not imply that the water is not hot.

The analogy between presuppositions and implicatures is not just imperfect: on closer inspection, it breaks down completely. There is no projection problem for conversational implicatures as there is a projection problem for presuppositions, and it isn't hard to see why, once the distinction between the notions of implicature and presupposition is appreciated. A conversational implicature is a kind of bonus inference, which is not available until it has been made out what the speaker is saying. When it is said that (38a) implicates (38b), for example, the idea is that on the basis of the truth-conditional content of (38a), one is entitled to infer that the speaker believes that (38b). This inference expands upon what the speaker commits himself to by tokening this sentence. Presuppositions are so-called because they are taken to be given in the discourse context. Logically speaking, what a speaker presupposes must be determined before it can be decided what he is saying, and what he is saying must be determined before it can be decided what he is implicating. There is no way presupposition projection could be explained in terms of conversational implicature.

## 1.4 Projection vs. scope

Informally, presuppositions may be described as interpretative elements that prefer to take wide scope. This is a sloppy way of speaking because presupposition projection has nothing to do with scope as grammarians employ that notion. Presupposition projection and scope taking are phenomena that must be clearly distinguished, although they interact in various ways. In this section I will concentrate on the first point, and attempt to show that the two phenomena are disjoint. I will touch only briefly on the interaction between presupposition projection and the interpretation of scope-bearing expressions, since this is one of the leitmotives of the second half of this book.[14]

The idea that presuppositional data might at least partly be explained in terms of scope has attracted a great many people. Most importantly, adherents of Russell's analysis of definite descriptions (which, after some wilderness years, seems to be enjoying a spectacular revival) in effect adopt a scope analysis of what on the presuppositional view is just one of many presupposition triggers. In comparison, treatments in terms of scope of other presuppositional expressions are few and far between. Incidents that I know of are a scope analysis of factive verbs proposed by Delacruz (1976), a similar analysis considered by Grice (1981), and Heim's (1992) tentative suggestion that aspectual verbs be treated in terms of scope.

To convey something of the attraction of this approach, let us quickly recapitulate Russell's (1905) famous theory of descriptions, as applied to his equally famous examples (49a, b).

(49)     a.    The King of France is bald.

         b.    The King of France is not bald.

Adopting a restricted-quantifier notation which I take to be self-explanatory, Russell's proposal is that the definite article is in fact a quantifier, which is defined as follows:

(50)     [the x: Px] Qx $=_{def}$ [some x: Px](Qx & [all y: Py](x = y))

With the help of this quantifier, (49a) is analysed as follows:

---

[14] Caveat: in this section I will use the notion of scope in the grammarian's sense. That is to say, scope is something which scope-bearing expressions have, and the scope of an expression is determined, roughly speaking, by moving it about at some level of analysis. Fairly uncontroversial examples are quantifying determiners, negation, modals, etcetera. The claim that presuppositions aren't scope-bearing elements in this sense does not to preclude the possibility that they interact with scope-bearing elements in some way or other. (Cf. the treatment of indefinites in DRT.)

(51)     [the x: King of France x](bald x)

This is true iff there is one and only one King of France, and he is bald. For our current purposes, the most interesting part of Russell's proposal is his analysis of (49b). Russell observes that this sentence has two possible readings. Presupposition theorists such as myself would say that this sentence can be construed with or without the presupposition that there is a King of France. But not Russell. He exploits the fact that according to his theory of descriptions, the definite NP is a quantified expression, which can have either wide or narrow scope with respect to the negation operator:

(52)     a.     [the x: King of France x]¬(bald x)

         b.     ¬[the x: King of France x](bald x)

(52a) represents the more likely reading of (49b), according to which this sentence says that there is a King of France, who is not bald. (52b) represents what is sometimes called the 'denial' reading of (49b), which merely says that it is not the case that there is a King of France who is bald. If it turns out that there is no King of France, (52b) is true and (52a) is false.

   This is certainly an elegant analysis, which moreover can be extended in a number of ways,[15] and it is a pity that something so beautiful could be so wrong. My reasons for maintaining that it is wrong are mostly the familiar ones, which were brought up by Strawson (1950). In particular, I have yet to see a convincing reply to Strawson's observation that a definite NP need not refer uniquely in order to be used felicitously. But this is a discussion that I don't want to get into here, because the current topic is the distinction between presupposition projection and scope, and in this connection the main problem with Russell's theory of descriptions is that it doesn't scale up properly, as an engineer would say. That is, if we want to extend the Russellian analysis to presuppositional expressions other than definite NPs, we are in deep trouble. To begin with, consider the following example:

(53)     a.     Many girls flirted with some teachers.

         b.     [many x: girl x][some y: teacher y](x flirted with y)

         c.     [some y: teacher y][many x: girl x](x flirted with y)

It is natural to give wide scope to *many girls*, as in (53b), and possible to give *some teachers* wide scope, as in (53c). But in either case, *many girls* (or *some teachers*) may be read as presupposing that there are or were girls (teachers). In other words, supposing that (53b) is the most likely analysis of (53a), what we need to account for is that the most natural paraphrase of this sentence is

---

[15] In particular, it can be extended to deal with certain varieties of anaphora. For some of the more recent developments, see Bach (1987), Heim (1990), and Neale (1990).

something like the following: 'There were girls and teachers and many of the former flirted with some of the latter.' If we want to account for this in terms of scope, we will have to assume, in effect, that the nouns *girls* and *teachers* can be given wide scope, along the following lines (here the capital variables range over sets or groups):[16]

(54)    [the X: girls X][the Y: teachers Y][many x: x ∈ X][some y: y ∈ Y]
         (x flirted with y)

This proposal may be supported by the observation that, on this reading, (53a) can be paraphrased with two explicit partitives:

(55)    Many of the girls flirted with some of the teachers.

But now consider the following example:

(56)    Everybody flirted with some teachers.

The most likely interpretation of this sentence is that every one of a contextually given group of individuals flirted with some teachers, so *everybody* in (56) gets the same sort of interpretation as *many girls* in (53a), and a Russellian analysis along the lines of (54) would have to look like this:

(57)    [the X: persons X][every x: x ∈ X][some y: teacher y]
         (x flirted with y)

Even if it can be shown that this is an adequate representation of the intended reading of (56), this analysis requires that *everybody* be decomposed into two parts, one of which is to be given wide scope; i.e. *everybody* must be read as 'every-one-of-the-people', and the second half of this representation must be fronted. Needless to say, this is not a very plausible procedure. I have no qualms about lexical decomposition *per se*, but if we must assume that rules of interpretation can extract parts from the semantic representation of a word, I begin to feel uncomfortable (cf. Geurts 1996b).

Let me try to bring the problem into clearer focus with the help of the following examples.

(58)    a.    Perhaps Fred has stopped dating Betty.

        b.    Perhaps Fred has started dating Betty.

Normally speaking, someone who utters (58a) presupposes that Fred dated Betty before a given time, while an utterance of (58b) will normally imply that Fred did not date Mary before that time. If we want to represent these

---

[16] The quantifier *the* employed in (54) is obviously not the same as the *the* defined in (50), which will only deal with definite NPs in the singular. See Neale (1990) for a Russellian analysis of

readings in terms of scope, we have to make some pretty fancy assumptions, but I am willing to grant practically anything if only I can get my main point across. Here, then, is how (58a) might be represented:

(59)     [the e: Fred dated Betty during e] ◇(Fred stopped e)

The intended interpretation of (59) is that there is a unique event e during which Fred dated Betty, and that, possibly, Fred discontinued e. Supposing that this is right, how did we get from (58a) to (59)? One answer to this question is that we construed the gerund *dating Betty* as a definite, and gave it wide scope. This cannot be right, however, because this procedure would give the wrong result for (58b):

(60)     [the e: Fred dated Betty during e] ◇(Fred started e)

This says that Fred used to date Betty, whereas it should say the opposite. It seems, therefore, that we have to claim that in both cases part of the interpretation of the whole predicate is extracted to give it wide scope. So *stopped dating Betty* is parsed as 'dated Betty and stopped doing so', and *started dating Betty* is parsed as 'did not date Betty but started doing so', and in both cases the first half of the analysis is construed as a definite expression, which is given wide scope. I don't think it will be necessary to spell out why this line of analysis is not very promising.

   Thus far it would seem that a scope analysis of projection phenomena is not impossible to carry through, although it will require a number of unlikely assumptions. But there are also data that are definitely outside the purview of a scope analysis.

(61)     a.   Most professors sold their Coca Cola shares.

         b.   Theo usually stops drinking before he gets sick.

         c.   Everyone should leave his camera at the information desk.

What these sentences have in common is that, on their most natural readings, a presuppositional expression in the predicate somehow restricts the domain of a quantifier. Thus, (61a) will be read as saying that most of the professors *who owned Coca Cola shares* sold them; (61b) says that usually, *if Theo drinks*, he stops before he gets sick; and the most likely interpretation of (61c) is that everyone *who has a camera* should leave it at the information desk. Apparently, the sources of these restrictions are the presuppositions triggered by *their Coca Cola shares*, *stops drinking*, and *his camera*, respectively. These readings cannot be accounted for in terms of scope. In (61a), for instance, there are two scope-bearing expressions (if Russell is right), one of which contains a pronoun which is bound by the first, so the only reading a Russellian theory will deliver for this sentence is something like the following:

(62)      [most x: professor x][the Y: Y are Coca Cola shares owned by x]
            (x sold Y)

This may or may not be a possible reading of (61a), but it certainly does not represent the interpretation we want to have, since (61a) might be true, while (62) would be false, if it weren't the case that most professors owned Coca Cola shares. The same holds, mutatis mutandis, for (61b, c).

These observations will suffice to discredit the notion that projection phenomena can be accounted for in terms of scope. The reason for this is that presupposition projection is a pragmatic phenomenon, whereas scope taking is much more of a surface phenomenon. If we speak about an expression α in terms of scope, we are referring to some α-sized unit, i.e. α itself or some semantic entitity corresponding with α, such as α's correlate at LF or its interpretation in a given model. Although in some frameworks this is only a metaphor, the guiding intuition is that α's scope is determined by moving α about. If, on the other hand, we speak of a presupposition triggered by α, the metaphor is a completely different one, the idea being, rather, that α requires us to check that some piece of information is given. This distinction is perhaps most evident in expressions that have scope and trigger presuppositions at the same time, like quantified NPs. An NP such as *most flowers*, for example, takes scope with respect to other parts of the sentence in which it occurs, while at the same time it triggers a presupposition, viz. that some collection of flowers is contextually given.

The distinction between scope taking and presupposition projection is obscured by the unfortunate circumstance that definite NPs, which have always been the presuppositional expressions *par excellence*, happen to have the special property that their descriptive contents coincide with the presuppositions they induce. On the presuppositional account of definitess that I advocate, a definite like *the banana* induces the presupposition that some banana is given in the discourse context, and to a first approximation that is all there is to say about the content of *the banana*. In general, the content of an expression divides into an asserted and a presupposed part. Definites are special in that the asserted part is empty. Consequently, unless the presupposition triggered by a definite NP α is blocked, which will rarely happen, it will seem as if α itself had taken wide scope, but this is an illusion, as we have seen, because definite NPs aren't scope-bearing expressions.

Although it wasn't the main purpose of this section to argue against Russell's theory of descriptions, I would like to conclude with a remark on this theory's remarkable staying power. I will make my point by way of a simile. My little niece has developed her own theory about the evolution of forefingers.[17] It is not a very complex theory; in fact I can summarize it in seven words: forefingers have evolved for pointing at things. My niece

plural definites.

actually has a number of cogent arguments for this theory, but I will spare you the details, because it will be evident already that it is not a good theory at all. To begin with, it is strange that someone should even consider developing such a theory in the first place. The evolution of fingers would be a plausible subject; the evolution of forefingers is obviously not. And to make things even worse, it turns out that there is no way this theory could be extended so as to deal with the evolution of the other fingers. In brief: without looking at the details we can tell beforehand that the theory is ad hoc.

Russell's theory of descriptions is like my niece's theory about forefingers. It confines its attention to a small portion of a broad class of related facts, and this limitation is essential to whatever plausibility it may seem to have. If my niece is convinced that her theory of forefingers is a sound one, it is only because she blithely ignores the other fingers. Similarly, one can only defend Russell's theory if he closes his eyes to the fact that definite NPs make up only a fraction of a large complex of related phenomena, which the theory cannot account for.

Of course, Russellians are bound to deny that there are non-trivial similarities between definite NPs and factives, clefts, aspectual verbs, and so on. A thoroughbred Russellian will rather say something like this:

> A great range of disparate and unrelated phenomena has been dubbed 'presuppositional' over the years, but [...] it seems highly implausible that any theoretically important notion will do justice to the full range of data that semanticists professing an interest in 'presupposition' seek to explain. (Neale 1990: 54)

I have been at pains to show that there is wide range of empirical phenomena which cry out for a common explanation. This is what presupposition theorists call the projection problem. One can of course choose to ignore this problem or deny that it exists, as Neale does, but to do so is like saying that there is no theoretically important notion that can explain why fingers have evolved, because your niece has developed such a neat account of the evolution of forefingers.

## 1.5  Where do they come from?

In the following chapters the focus will on the problem of presupposition projection, and it will be taken for granted that factive verbs presuppose their clausal complements, that a definite NP of the form 'the N' triggers the presupposition that some N is contextually given, and so on. In short, it will be assumed that presuppositions are triggered by certain lexical items and

syntactic constructions. In this respect I will continue a long tradition, because presupposition theorists have mostly been interested in two problems only: presupposition projection and presupposition failure. Both problems require for their formulation that presuppositions come about one way or another, and nearly everybody has been contented to take this for granted. Indeed, it has been doubted that the problem where presuppositions come from deserves serious attention at all. Gazdar (1979: 126), for example, considers it to be 'a theoretically trivial task' to specify a function that takes words and syntactic types into presuppositions. I maintain that Gazdar is wrong about this, and that the question of where and how presuppositions originate merits serious investigation. And although I don't have anything like a complete answer to this question, the least I can do is try and show that the question is a good and hard one.

Although there has been little interest in the question where presuppositions come from, there does seem to be a tacit consensus on how the question must be answered. It is that presuppositions are induced grammatically (and therefore, conventionally) by certain lexical elements and syntactic constructions.[18] I can think of only one case for which this view is really plausible. It seems to me that it can hardly be doubted that the definite article is a grammatical device for marking presuppositions. Definite articles are flags for signalling that certain pieces of information are taken for granted, but already the fact that many languages do not have definite articles should make us wary of the conventional wisdom that presuppositions always originate in the grammar. And indeed, as soon as we turn to presupposition-inducing expressions other than the definite article, this piece of linguistic lore starts to crumble.

Let us look at some examples. So-called 'aspectual' verbs like *begin* or *stop* are presupposition-inducing expressions:

(63)    a.    Our neighbour has stopped growing tulips.

        b.    Our neighbour has begun to grow tulips.

        c.    Our neighbour continues to grow tulips.

(63a) and (63c) presuppose that our neighbour used to grow tulips, whereas (63b) presupposes that he did not. Apparently, these presuppositions originate in the aspectual verbs *stop*, *begin*, and *continue*, and it is easy to stipulate that, for example, *stop* triggers the presupposition that the state described by its clausal complement held before the utterance's reference time. However, if we should simply encode this observation into the lexicon we would miss an obvious generalization. Each of the examples in (63)

---

17  Actually, I have *two* little nieces; but never mind.

describes a transition from one state to another: from growing tulips to not growing tulips (63a), from not growing tulips to growing tulips (63b), and from growing tulips to growing even more tulips (63c). In each case it is the *initial* state whose existence is presupposed, not the ensuing one, and this observation turns out to apply to all transition verbs (apart from those in (63) there are: *discontinue*, *go on*, *persist in*, and so on) as well as to particles like *still*, *already*, *anymore*, and so on. Of course, this regularity could be captured by devising a rule that states that all these items have this property, instead of stipulating it for each separately,[19] but one suspects that this rule should somehow follow from some semantic and/or pragmatic property that these expressions have in common.

It should be noted that, intuitively speaking, it is eminently *reasonable* that aspectual verbs should have the presuppositions that they in fact have. It seems unlikely that we should ever come across a verb that meant exactly the same thing as *begin* with the only difference that it presupposes what *begin* asserts and vice versa. And I believe that it is intuitively clear why this should be so. It is because interlocutors are more interested in where the story is leading to than where it came from, and therefore tend to take the past as given. If a change is described, the initial state is in a sense backgrounded because we are more interested in the present and the future. And this backgrounding corresponds with presupposition.

Before I try to draw a moral from this observation, let me introduce a second example. Factive predicates presuppose their complements. Compare:

(64)    It is $\left\{ \begin{array}{c} \text{nice} \\ \text{tragic} \\ \text{disgusting} \\ \text{odd} \end{array} \right\}$ that Fred is reading Kropotkin.

(65)    It is $\left\{ \begin{array}{c} \text{likely} \\ \text{possible} \\ \text{obvious} \\ \text{true} \end{array} \right\}$ that Fred is reading Kropotkin.

The standard tests show that the factive predicates in (64) trigger the presupposition that Fred is reading Kropotkin, whereas the predicates in (65) don't. Why is that? To begin with, note that, whereas all variants in (64) entail that Fred is reading Kropotkin, not all variants of (65) do. If it is obvious or true that Fred is reading Kropotkin, then Fred must be reading

Kropotkin; but this does not follow if is just likely or possible. Apparently, a predicate can entail its complement and not be factive. A more systematic difference between (64) and (65) is that the former are paraphrasable with 'the fact that', while the latter must rather be paraphrased with, say, 'the story that':

(66)    The $\left\{\begin{array}{c} \text{fact} \\ \text{?story} \end{array}\right\}$ that Fred is reading Kropotkin is $\left\{\begin{array}{c} \text{nice} \\ \text{tragic} \\ \text{disgusting} \\ \text{odd} \end{array}\right\}$.

(67)   The $\left\{\begin{array}{c} \text{*fact} \\ \text{story} \end{array}\right\}$ that Fred is reading Kropotkin is $\left\{\begin{array}{c} \text{likely to be true} \\ \text{possibly true} \\ \text{obviously true} \\ \text{true} \end{array}\right\}$.

(The question mark in (66) is not to imply that this variant is ill-formed but that it is inadequate as a paraphrase of (64).) This contrast does not yet explain the difference between the factive predicates in (64) and their non-factive counterparts in (65), however. Even if it is true that factives take facts as arguments, while non-factives don't, it doesn't follow from this that factives are presupposition inducers (there is no a priori reason why facts should be presupposed). Furthermore, this observation only reformulates the original problem: *Why* is it that certain predicates take facts as arguments?

The distinctive property of the non-factives in (65) is that they comment on the alethic or epistemic status of their arguments, and this is why they cannot take their truth for granted. If someone *asserts* the truth of a proposition, he cannot simultaneously presuppose that it is true, and if something is claimed to be likely or possible, it cannot be presupposed, either. But even if this explains why the examples in (65) do not presuppose that Fred is reading Kropotkin, it does not explain why the examples in (64) do. What I want to suggest, if only tentatively, is that the factives in (64) induce the presupposition that Fred is reading Kropotkin because there is no reason why they should not. A speaker who asserts that it is disgusting that Fred is reading Kropotkin conveys the two-part message (i) that Fred is reading Kropotkin, and (ii) that this is disgusting. The second part contains the main point the speaker has to make; the first part is a prerequisite for the first, but it is of secondary importance. In this sense, the first part is backgrounded — and therefore it is presupposed.

Generalizing from these examples, my suggestion is the following. The content of an utterance is complex, not only at sentence level but also below that; even the content of a single word will rarely be a simple matter. In view

of this complexity, it is natural that the interlocutors will concentrate their attention on selected parts of the content conveyed by an utterance; the rest is of secondary importance, it is backgrounded. There may be many factors that can influence this selection process, but once the focal points have been identified, what remains tends to be presupposed.

This picture is admittedly vague, and it certainly does not deserves to be called a theory, but it is not entirely without substance. The main idea is that, although there may be many different reasons for backgrounding part of the content of an expression, what is judged to be of secondary interest will be presupposed if it can be presupposed. This allows us make some sense of the bewildering variety of presupposition triggers. It may be that, say, factives and aspectual verbs have little in common. But both types of expressions focus on certain parts of the information they contain, and allow the remainder to be presupposed, and if I am right, this is why it will be presupposed. It will be evident that there is much more to say about this subject. Unfortunately, however, I don't have much more to say about it.

## 1.6  A note on presupposition failure

Of the three main problems that I mentioned at the beginning of this chapter, the problem of presupposition failure is the oldest, and in my opinion it is also the least interesting. What happens if a presupposition turns out to be false? Suppose that Betty isn't married. Is (68) then false or just meaningless? Will a speaker who uttered this sentence have *said* something in any non-trivial sense of the word, or not?

(68)     Betty's husband is a teetotaler.

It is over questions like these that the famous battles between Fregeans, Russellians, and Strawsonians have raged, but despite the controversy caused by the issue of presupposition failure, I feel that its importance has been overrated. To begin with, taking a purely observational stance, it is by no means clear what the issue is supposed to be, with informants being inconsistent over what the consequences of presupposition failure are. Given that Betty isn't married, some native speakers of English will say that (68) is false, while others will prefer to say that it doesn't make any sense. The empirical bottom line is that the opinions about the effects of presupposition failure are divided, and to the extent that a consensus can be attained, it appears that speakers' intuitions are dependent on contextual factors of a such a kind that only further doubt is cast upon the presumption that there is a substantial issue. At any rate, this is the moral I draw from Strawson's

observation that speakers' intuitions about presupposition failure are partly
determined by what they take to be the topic of discourse:

> Confronted with the classical example, 'The king of France is bald',
> we may well feel it natural to say, straight off, that the question
> whether the statement is true or false doesn't arise because there is
> no king of France. But suppose the statement occurring in the
> context of a set of answers to the question: 'What examples, if any,
> are there of famous contemporary figures who are *bald*?' (Strawson
> 1964: 113)

In the latter case, Strawson notes, there is a strong inclination to say that the
answer is false, rather than meaningless. In my opinion, these observations
are clearly correct, and they suggest rather forcefully that, although
speakers' intuitions about presupposition failure are not entirely devoid of
interest, they shouldn't be allowed to play a central role in a theory of
presupposition.[20] As Stalnaker puts it:

> I do not think any of us have very clear intuitions about the truth
> values of statements which have false presuppositions, and so I do
> not think that the truth value, or lack of it, of such statements can
> be data against which to test competing generalizations. (Stalnaker
> 1973: 454)

I will have nothing of interest to contribute to the ongoing debate about
presupposition failure.

---

[18] This is very explicit in all of Karttunen's and Gazdar's work, for example, and in van der
Sandt's earlier contributions to the literature (such as van der Sandt 1987, 1988). More recent
publications tend to be less explicit on this point.
[19] The possibility of such a rule is mentioned by Gazdar (1979: 128).

CHAPTER 2

# The binding theory

Although discourse representation theory aspires to be a general theory of interpretation, its main applications have been in the areas of anaphoric and temporal reference. It is in particular the first area that I am interested in, because I maintain, following van der Sandt, that the DRT account of anaphora can be generalized to a theory of presupposition projection.

  In this chapter, I present a theory of presupposition projection which is an extension of DRT, and therefore I start off with a brief outline of the DRT framework (§ 2.1). In the next two sections, it is shown how DRT can be extended to a theory of presupposition projection (§§ 2.2 and 2.3). The result of this extension will be the binding theory, which is so-called because it views presuppositions as elements that want to be bound in more or less the same sense in which anaphoric elements want to be bound. These two sections are the centrepiece of this book, for although the theory presented there will be refined in the second half of this chapter, they introduce the main ingredients for my account of presupposition projection in intensional contexts (Chapters 5 and 6), and my analysis of names (Chapter 7). Indeed, the binding theory as presented here will not be modified in later chapters, as I will argue that, together with independently motivated treatments of modal expressions and attitude verbs, it explains the behaviour of presuppositions in modal contexts and attitude reports. Nor will my theory of names require any modifications or extensions of the binding theory as presented in this chapter. Having illustrated and refined the theory with the help of a few small case studies in § 2.4, the chapter closes with some further remarks on the key concepts of the binding theory, and their relation to the traditional notion of anaphora.

## 2.1 Discourse representation theory

As its name already indicates, discourse representation theory is concerned with the interpretation of discourses, not just sentences.[1] As a discourse

---

[1] DRT was originally proposed by Kamp (1981). What I will sometimes refer to as the standard version of the theory is Kamp and Reyle's (1993).

unfolds, interlocutors incrementally construct a model of what has been said. These models are called 'discourse representation structures' (DRSs). DRSs may be viewed as representations of commitment slates in the sense of Hamblin (1971, 1987: 229–232). A speaker who utters a sentence φ makes a commitment of sorts. He need not actually believe that φ is true, but he is committed to act, at least for the time being, as if he believed φ to be true. Communication doesn't break down when I say something I don't believe in, but when I assert something and deny its truth immediately afterwards, the communication process is in an impasse. We may imagine that each speaker has a slate on which his conversational commitments are recorded, and the interlocutors maintain representations of each other's commitment slates. In DRT such representations are taken to be DRSs.

A DRS consists of two parts: a universe of so-called 'reference markers', which represent the objects under discussion, and a set of DRS-conditions which encode the information that has accumulated on these reference markers. The following DRS represents the information that there are two individuals, one of which is a policeman, the other a squirrel, and that the former chased the latter:

(1)      [x, y: policeman x, squirrel y, x chased y]

The universe of this DRS contains two reference markers, x and y, and its condition set is {policeman x, squirrel y, x chased y}.

A DRS like the one in (1) can be given a straightforward model-theoretic interpretation. In DRT this is usually done by defining what it means for a function to be a verifying embedding of a DRS into a given model. A function $f$ embeds (1) into a model M iff the domain of $f$ is {x, y}, $f(x)$ is a policeman in M, $f(y)$ is a squirrel, and $f(x)$ chased $f(y)$. Hence, embedding functions are partial variable assignments.

Meanwhile it will have become clear that the DRS in (1) reflects the intuitive meaning of:

(2)      A policeman chased a squirrel.

and indeed, it is claimed that, in the absence of any information about the context in which this sentence is uttered, the semantic representation of (2) is (1). So the indefinite subject and object NPs in this sentence cause the introduction of two new reference markers, x and y, and contribute the information that x is a policeman and y a squirrel, and the verb contributes the information that x chased y.

If a discourse begins with an utterance of (2), the DRS in (1) is constructed, and this DRS forms the background against which the following utterance is interpreted. Suppose now that (2) is followed by a token of (3a):

(3)      a.    He caught it.

b.    [v̲, w̲: v caught w]

(3b) is the DRS that reflects the semantic content of (3a) before the pronouns are resolved. In this DRS, the anaphoric pronouns *he* and *it* in (3a) are represented by the reference markers v and w, respectively. These reference markers are underlined to indicate that they require an antecedent. Since (3a) is uttered in the context of (1), the next step in the interpretation of this sentence is to merge the DRS in (3b) with that in (1). The result of this merging operation is (4a):

(4)    a.    [x, y, v̲, w̲: policeman x, squirrel y, x chased y, v caught w]

       b.    [x, y, v, w: v = x, w = y,
                policeman x, squirrel y, x chased y, v caught w]

       c.    [x, y: policeman x, squirrel y, x chased y, x caught y]

Since (3a) is immediately preceded by (2), the subject pronoun *he* is probably intended to be co-referential with *a policeman*, while *it* should link up to *a squirrel*. At DRS level, this is represented by equating the reference markers v and w with x and y, respectively. These operations yield (4b), which is equivalent to (4c). Either DRS is embeddable in a model M iff M features a policeman who chased and caught a squirrel.

   We have said that the contribution of (3a) to the context represented by (1) is obtained in two steps. In the first step, a DRS is constructed which is determined compositionally — i.e., it is a function of the lexical contents and the syntactical constructions that enter into the grammatical analysis of the sentence. In the second step, this sentence DRS is merged with the DRS that represents the prior discourse, and anaphoric references are resolved. This two-stage procedure is also employed by Asher (1993), but the original version of DRT (Kamp 1981, Kamp and Reyle 1993) is monostratal in the sense that one set of rules takes care of both tasks at once. I prefer the two-stage model for the following reasons. First, as we will be mostly concerned with pragmatic aspects of interpretation, it is convenient that we don't have to worry about the grammatical details of our analyses, which as a rule will be uncontroversial anyhow (the grammatical details, I mean). Secondly, I believe that as a matter of fact the division between grammar and pragmatics which the two-stage model requires is empirically justified, and that a monostratal account blurs a natural boundary line. This complaint obviously cannot be justified on a priori grounds. It requires that we consider in some detail a sizeable amount of data. But that is precisely what I intend to do in the following. So this book may be viewed, inter alia, as an extended argument in favour of the grammar/pragmatics divide as suggested by the two-stage model.

   The term 'two-stage model' is not to suggest that the grammatical analysis of a sentence must be completed before pragmatic processes come into play.

Although in the following I will always start out from sentence-sized semantic representations, this is only because I want to focus on pragmatic aspects of interpretation. But it will be clear that in reality semantic and pragmatic processes are interleaved. In this respect the standard DRT model is more realistic than the one adopted here, although I take it that the two-stage model could be reformulated so as to allow the pragmatic processes to follow the grammar at each step. However, unlike standard DRT, this reformulated version would still make a clean distinction between linguistic and pragmatic determinants of meaning. To give an example, Kamp and Reyle (1983: 122) have a single rule for interpreting pronouns, which is triggered whenever a parse tree contains a pronominal element. The rule says that for each pronoun we must select a new reference marker, link it to a suitable antecedent, and update the DRS so as to record these changes. In my version of DRT there is no such rule. Its duties are divided between two separate mechanisms. On the one hand, pronominal lexemes introduce new reference markers and mark them as anaphoric. On the other hand, there is a completely general pragmatic mechanism for dealing with presuppositions, which handles anaphoric reference markers, too. So the main difference between my version of DRT and Kamp and Reyle's is that mine has a separate pragmatic component which takes care of certain tasks that in the Kamp and Reyle system are dealt with, whenever the need arises, by the individual construction rules.

Thus far, we have only considered DRSs with simple conditions, but in order to account for negated and conditional sentences, say, complex conditions are required.

(5)      a.      Wilbur doesn't have a radio.
         b.      $[_1$ x: Wilbur x, $\neg[_2$ y: radio y, x owns y]]

(5b) is the sentence DRS corresponding to (5a). This DRS contains a condition which consists of a DRS prefixed by a negation sign. It is sometimes convenient to adorn DRSs with numerical labels, as I have done in (5b), and refer to these labeled DRSs with names like '$(5b_1)$', '$(5b_2)$', and so on. In general, labeling of DRSs will be top-down and left-to-right, so the main (or principal) DRS will always be number one.

A function $f$ embeds $(5b_1)$ in a given model M iff $f$ maps x onto an individual in M which 'is a Wilbur', i.e. which is called 'Wilbur', and $f$ cannot be extended to a function $g$ which embeds $(5b_2)$ — that is to say, no such $g$ should map y onto a radio that Wilbur owns. Proper names are treated here as ordinary definite NPs, which are accounted for by the theory of presupposition. This treatment will be defended in Chapter 7.

$(5b_2)$ contains a token of the reference marker x which is introduced externally, in the DRS in which $(5b_2)$ is embedded, i.e. $(5b_1)$. Apart from that, $(5b_2)$ also introduces a reference marker of its own, i.e. y, which is

correlated with the indefinite NP *a radio*, and whose scope is delimited by
$(5b_2)$. Consequently, it doesn't make sense to refer to y outside of $(5b_2)$. In
DRT, this is taken to explain why the 'lifespan' of the individual introduced
by the indefinite NP in (5a) is delimited by the scope of the negation
operator. If (5a) were followed by (6a), for example, the pronoun could not
be linked to the indefinite:

(6)  a.  It's a Sony.

  b.  $[\underline{z}: \text{Sony z}]$

  c.  $[_1 \text{ x}, \underline{z}: \text{Wilbur x}, \neg[_2 \text{ y}: \text{radio y}, \text{x owns y}], \text{Sony z}]$

If we merge (5b) and (6b), which is the sentence DRS corresponding with
(6a), we obtain (6c). In this representation, the reference marker z does not
have access to y, because y is introduced in a DRS that is not accessible to the
DRS in which z is introduced, and therefore it is not possible to bind z to y.
In other words, if (6a) is preceded by (5a), the pronoun cannot be
anaphorically linked to the indefinite NP. This prediction appears to be
correct.

Accessibility is primarily a relation between DRSs; derivatively, it is also a
relation between reference markers. $(6c_1)$ is accessible to $(6c_2)$, but not the
other way round, and therefore the reference markers introduced in $(6c_1)$,
i.e. x and z, are accessible from $(6c_2)$, but conversely, if we are in $(6c_1)$ we
have no access to $(6c_2)$ and its reference markers, i.e. y. Thus in (6c) anaphora
is not possible because y is not accessible to z. In (4a), by contrast, anaphora
is possible, because x and y are accessible to v and w (the accessibility
relation being reflexive). The notion of accessibility is crucial to DRT's
account of anaphora, and it is important to bear in mind that it is tied up with
the semantics of the DRS language. $(6c_1)$ is accessible to $(6c_2)$ because every
embedding function that must be considered for $(6c_2)$ is an extension of an
embedding function for $(6c_1)$, and it is for this reason that every reference
marker in $(6c_1)$ is also defined in $(6c_2)$. The converse, however, does not
hold.

Like negated sentences, conditionals give rise to complex DRS-conditions,
too. (7) gives an example:

(7)  a.  If Leo chased a squirrel, he caught it.

  b.  $[_1 \text{ x}: \text{Leo x}, [_2 \text{ y}: \text{squirrel y}, \text{x chased y}] \Rightarrow [_3 \underline{v}, \underline{w}: \text{v caught w}]]$

  c.  $[_1 \text{ x}, \text{v}: \text{v} = \text{x}, \text{Leo x},$
    $[_2 \text{ y}, \text{w}: \text{w} = \text{y}, \text{squirrel y}, \text{x chased y}] \Rightarrow [_3 : \text{v caught w}]]$

  d.  $[_1 \text{ x}: \text{Leo x}, [_2 \text{ y}: \text{squirrel y}, \text{x chased y}] \Rightarrow [_3 : \text{x caught y}]]$

(7b) is the sentence DRS which corresponds to (7a), and assuming for
convenience that this sentence is uttered in an empty context, it is also the

initial DRS of the discourse. The complex condition in this structure is interpreted as follows: if $f$ is to be an embedding function for $(b_1)$, then it must map x onto an individual called 'Leo', and every extension of $f$ which embeds $(b_2)$ must itself be extendable to a verifying embedding of $(b_3)$. It follows from this that $(b_1)$ is accessible to $(b_2)$, which in its turn is accessible to $(b_3)$, and therefore v may be linked up to x (accessibility being a transitive relation) and w to y. The result is (7c), which is equivalent to (7d), and the meaning of either of these is the same as that of the following formula in predicate logic:

(8)     $\exists$x[Leo x $\wedge$ $\forall$y[[squirrel y $\wedge$ x chased y] $\rightarrow$ x caught y]]

To round out our inventory of complex DRS-conditions, we turn to disjunctive sentences.

(9)     a.     Leo chased and caught a squirrel. He either handed it over to a policeman or he took it home with him.

        b.     $[_1$ x, y: Leo x, squirrel y, x chased y, x caught y,
                $[_2$ z, v̠: policeman z, x handed over v to z] $\vee$
                $[_3$ w̠: x took w home]]

        c.     $[_1$ x, y, v, w: v = y, w = y,
                Leo x, squirrel y, x chased y, x caught y,
                $[_2$ z: policeman z, x handed over v to z] $\vee$
                $[_3$ : x took w home]]

        d.     $[_1$ x, y: Leo x, squirrel y, x chased y, x caught y,
                $[_2$ z: policeman z, x handed over y to z] $\vee$
                $[_3$ : x took y home]]

We start out from the DRS in (9b), in which the pronoun *he* has already been resolved. Note that the conjunction in the first sentence of (9a) is represented in (9b) simply by merging the DRSs correlated with the two conjuncts. The interpretation of disjunctive DRS-conditions is straightforward: an embedding function verifies a condition of the form '$\varphi$ $\vee$ $\psi$' iff it either verifies $\varphi$ or $\psi$. It follows from this that $(9b_1)$ is accessible to $(9b_2)$ and $(9b_3)$, and therefore the reference markers v and w can link up to their intended antecedent, i.e. y. The resulting DRS is (9c), which is equivalent to (9d).

We have seen that, in order to embed a condition of the form '$\varphi$ $\Rightarrow$ $\psi$' into a model, each function which embeds $\varphi$ must have an extension which embeds $\psi$ as well. As a consequence, $\varphi$ is accessible to $\psi$, and anaphora from $\psi$ into $\varphi$ becomes possible. The members of a disjunction, however, are interpreted in parallel, so neither disjunct is accessible to the other. It is predicted, therefore, that anaphora from one disjunct into the other is impossible. This prediction is problematic in view of the following type of example, first discussed in modern times by Geach (1962):

(10)    a.    Either Wilbur doesn't have a radio or he keeps it hidden somewhere.

b.    $[_2$ x: Wilbur x,
$[_2 : \neg[_3$ y: radio y, x owns y]] $\vee$ $[_4$ z: x keep z hidden]]

The reference marker z in $(10b_4)$ correlates with the anaphor *it* in the second disjunct of (10a). Intuitively, this anaphor has something to do with the indefinite *a radio* in the first disjunct, but it is clear that it cannot be bound by it. In DRT terms, $(10b_3)$ is not to accessible $(10b_4)$ (nor vice versa, for that matter), and therefore z cannot be bound to y. This is one of the problems that will be addressed in the following sections.

We now proceed to give a more succinct formulation of the theory outlined in the foregoing, starting with a definition of the DRS language.

(11)    *DRSs and DRS-conditions*

a.    A DRS $\varphi$ is a pair $\langle U(\varphi), Con(\varphi) \rangle$, where $U(\varphi)$ is a set of reference markers, and $Con(\varphi)$ is a set of DRS-conditions.

b.    If P is an *n*-place predicate, and $u_1, ..., u_n$ are reference markers, then $Pu_1 ... u_n$ is a DRS-condition.

c.    If *u* and *v* are reference markers, then $u = v$ is a DRS-condition.

d.    If $\varphi$ and $\psi$ are DRSs, then $\neg\varphi$, $\varphi \vee \psi$, and $\varphi \Rightarrow \psi$ are DRS-conditions.

The notion of accessibility is defined in (12) (this particular formulation is inspired by Kadmon's 1990).

(12)    *Accessibility*

$\leq$ is the smallest preorder (transitive, reflexive) for which all of the following hold, for any $\varphi$, $\psi$, $\chi$:

a.    If $\neg\psi\ Con(\varphi)$, then $\varphi \leq \psi$

b.    If $\psi \vee \chi \in Con(\varphi)$, then $\varphi \leq \psi$ and $\varphi \leq \chi$

c.    If $\psi \Rightarrow \chi \in Con(\varphi)$, then $\varphi \leq \psi \leq \chi$

If $\varphi \leq \psi$, then we say that $\varphi$ is accessible to $\psi$. Since $\leq$ is a reflexive relation, each DRS is accessible to itself. Also, since $\leq$ is transitive, it is ensured, for example, that $(7d_1)$ is accessible to $(7d_3)$.

In terms of the notion of accessibility, we define in (13) which reference markers are accessible from a given DRS:

(13)    *Accessible domains*

$Acc(\varphi) = \cup\{U(\psi) \mid \psi \leq \varphi\}$

A reference marker $u$ is accessible from $\varphi$ iff $u \in U(\psi)$, for some $\psi$ that is accessible to $\varphi$ (which may be $\varphi$ itself). This notion is an important one, because the set of accessible reference markers of $\varphi$ is the set of possible antecedents for any anaphors occurring in $\varphi$.

One general constraint on the interpretation of an utterance is that it should yield DRSs that are proper (see § 2.3). The class of proper DRSs is defined as follows:

(14)    *Employment*

Emp($\varphi$) is the smallest set of reference markers for which the following hold:

a.    If $Pu_1 \ldots u_n \in \text{Con}(\varphi)$, then $\{u_1, \ldots, u_n\} \subseteq \text{Emp}(\varphi)$

b.    If $u = v \in \text{Con}(\varphi)$, then $\{u, v\} \subseteq \text{Emp}(\varphi)$

(15)    *Properness*

$\varphi$ is a proper DRS iff $\forall \psi \geq \varphi$: $\text{Emp}(\psi) \subseteq \text{Acc}(\psi)$

For example, the DRS [x: potato x, x loves y] is not a proper DRS, because it employs a reference marker, y, which is not in its accessible domain.

Taken together, DRSs and accessibility form the DRT counterpart to Stalnaker's notion of local (or intermediate) context, which was discussed in § 1.2. Stalnaker would say that, whenever a sentence of the form *If* S *then* S′ is uttered in a context c, the local context of S′ consists of the information in c plus the information provided by S. This translates into the DRT framework as follows: the context c is represented by the principal DRS $\varphi$, which contains a condition of the form $\psi \Rightarrow \psi'$, where $\psi$ and $\psi'$ correspond to S and S′, respectively. Talk of contexts thus gives way to talk of DRSs and accessibility, and instead of saying that the local context of S′ is c-plus-S, we now say that $\varphi$ and $\psi$ are accessible to $\psi'$.

The central concept in the semantics of DRT is that of an embedding function. An embedding function for a DRS $\varphi$ is an assignment of values to all and only the reference markers in $\text{Acc}(\varphi)$ which makes every condition in $\varphi$ come out true.[2] This is essentially what (16a) says. After the foregoing discussion, the remainder of this definition should not bring any surprises. The semantics in (16) is non-compositional because we must take into account the fact that a DRS may be interpreted in the context of larger DRS.

---

[2]    My notion of embedding function entails that no $f$ can be a verifying embedding for $\varphi$ unless $\text{dom}(f) = \text{Acc}(\varphi)$. This is more restrictive than the standard notion, which merely requires that $\text{Acc}(\varphi) \subseteq \text{dom}(f)$. Apart from the fact that I find my version more intuitive (why should $f$ be allowed to be defined for reference markers that don't occur in $\varphi$?), there are purely technical considerations that have prompted this restriction. These will become relevant when we extend the DRS language in Chapter 5.

The content of this larger DRS, if present, becomes essential at the point at which reference is made to Acc($\varphi$), in (16a). This context dependence could be encoded with the help of further subscripts, say, but the explicitness thus gained would be offset by a loss in transparency.

(16)     *DRS semantics (first version)*

Let M = $\langle$D, I$\rangle$ be a model, where D is a domain of individuals and I is an interpretation function. Let $f$ be a partial function from the set of reference markers into D.

a.     $M \vDash_f \varphi$ iff $\exists g \supseteq f$: $\text{dom}(g) = \text{Acc}(\varphi)$ and $\forall \psi \in \text{Con}(\varphi)$: $M \vDash_g \psi$

b.     $M \vDash_f Pu_1 \dots u_n$ iff $\langle f(u_1), \dots, f(u_n) \rangle \in I(P)$

c.     $M \vDash_f u = v$ iff $f(u) = f(v)$

d.     $M \vDash_f \neg\varphi$ iff $\neg\exists g \supseteq f$: $M \vDash_g \varphi$

e.     $M \vDash_f \varphi \vee \psi$ iff $\exists g \supseteq f$: $M \vDash_g \varphi$ or $M \vDash_g \psi$

f.     $M \vDash_f \varphi \Rightarrow \psi$ iff $\forall g \supseteq f$: if $M \vDash_g \varphi$ then $\exists h \supseteq g$: $M \vDash_h \psi$

'$M \vDash_f \varphi$' may be read as '$f$ embeds $\varphi$ into the model M'. Note the correspondence between (16d–f) on the one hand and (12a–c) on the other: if on the semantic side any embedding functions that we must consider for a DRS $\psi$ are extensions of an embedding for $\varphi$, then $\varphi$ is accessible to $\psi$. Thus the constraints on anaphora that DRT predicts mirror the interpretation of the DRS language.

In defining the notions of entailment and consistency, it is useful to have the following operation:

(17)     *DRS-merge*

If $K$ is a set of DRSs, then:
$\oplus K = \langle \cup\{U(\varphi) \mid \varphi \in K\}, \cup\{\text{Con}(\varphi) \mid \varphi \in K\} \rangle$

The merge of a set of DRSs simply is the pointwise union of their components. Entailment and consistency can now be defined as follows:

(18)     *Entailment*

A set of DRSs $K$ entails $\varphi$ iff $\forall M, f$:
if $M \vDash_f \oplus K$ then $\exists g \supseteq f$: $M \vDash_g \oplus(K \cup \{\varphi\})$
If $K = \{\psi\}$, and $K$ entails $\varphi$, then we will say that $\psi$ entails $\varphi$.

(19)     *Consistency*

A set of DRSs $K$ is consistent iff $\exists M, f$: $M \vDash_f \oplus K$

If $K = \{\varphi, \psi\}$, and $K$ is consistent, then we will say that $\varphi$ and $\psi$ are consistent with each other. If $K = \{\varphi\}$, and $K$ is consistent, then we will say that $\varphi$ is consistent.

The second of these definitions is straightforward. It says that a set of DRSs *K* is consistent iff a model and an embedding function can be found for $\oplus K$. The definition of entailment in (18) says that *K* entails φ iff, for every model M and every function *f* that embeds $\oplus K$ into M, we can find a function *g* which extends *f* and embeds the merge of *K* and φ into M. This definition (or something like it) is required in view of the DRS semantics in (16). The point is that a function can never embed a DRS into any model if the function is too big. For example, if *f* is a function which maps x onto a tomato and y onto a potato, then *f* can never be an embedding for [x: tomato x], because this DRS doesn't say anything about y. However, although no function which embeds [x, y: tomato x, potato y] into a given model can ever be extended to a function which embeds [x: tomato x], we do want the set {[x: tomato x], [y: potato y]} to entail [x: tomato x], and it is for this reason that we take the merge of the premise set and [x: tomato x], which in this case is [x, y: tomato x, potato y].

What I have presented in the foregoing is essentially the standard version of DRT. In the remainder of this section I will first reformulate the semantics given in (16), and then extend the DRS language with a possibility operator. The reason for doing this is that I want to make preparations for the extensions to DRT that will be proposed in Chapters 5 and 6.

In (16) the semantics of the DRS language is given by stating what it means for a function to embed a DRS in a model. That is, we have defined $M \vDash_f \varphi$, for any M, *f*, and φ. Given this relation, we might say that, relative to an embedding function *f*, each DRS denotes a set of functions *g* which extend *f* and which embed φ in M. For example, $[[x: \text{banana } x]]_{M,f}$ (the denotation of the DRS [x: banana x] in the model M relative to the function *f*) would be the set of functions which extend *f*, whose domain is {x}, and which map x onto a banana in M (this set may be empty, of course). In Chapter 5, I will define an extended version of DRT whose semantics will be given directly in terms of [.], and therefore I want to show here how the same can be done for the standard DRS language:

(20)    *DRS semantics (second version)*

Let M = ⟨D, I⟩ be a model, where D is some domain and I is an interpretation function. Let *f* be a partial function from the set of reference markers to D. $[\varphi]_{M,f}$ is either 1 or 0 if φ is a condition, and a (possibly empty) set of embedding functions if φ is a DRS.

a.    $[\varphi]_{M,f} = \{g \mid f \subseteq g$ and $\text{dom}(g) = \text{Acc}(\varphi)$ and $\forall \psi \in \text{Con}(\varphi): [\psi]_{M,g} = 1\}$

b.    $[Pu_1 ... u_n]_{M,f} = 1$ iff $\langle f(u_1), ..., f(u_n) \rangle \in I(P)$

c.    $[u = v]_{M,f} = 1$ iff $f(u) = f(v)$

d.   $[\![\neg\varphi]\!]_{M,f} = 1$ iff $[\![\varphi]\!]_{M,f} = \varnothing$

e.   $[\![\varphi \vee \psi]\!]_{M,f} = 1$ iff $[\![\varphi]\!]_{M,f} \cup [\![\psi]\!]_{M,f} \neq \varnothing$

f.   $[\![\varphi \Rightarrow \psi]\!]_{M,f} = 1$ iff $\forall g \in [\![\varphi]\!]_{M,f} : [\![\psi]\!]_{M,g} \neq \varnothing$

To see the connection between this semantics and the previous one, it may be helpful to note that $\vDash$ and $[\![.]\!]$ are interdefinable. On the one hand, $[\![.]\!]$ can be defined in terms of $\vDash$ :

(21)   a.   If $\varphi$ is a DRS, then $[\![\varphi]\!]_{M,f} =_{def} \{g \mid f \subseteq g$ and $\mathrm{dom}(g) = \mathrm{Acc}(\varphi)$ and $M \vDash_g \varphi\}$

       b.   If $\varphi$ is a DRS-condition, then $[\![\varphi]\!]_{M,f} = 1 =_{def} M \vDash_f \varphi$

On the other hand, if we have $[\![.]\!]$, we can define $\vDash$ thus:

(22)   a.   If $\varphi$ is a DRS, then $M \vDash_f \varphi =_{def} [\![\varphi]\!]_{M,f} \neq \varnothing$

       b.   If $\varphi$ is a DRS-condition, then $M \vDash_f \varphi =_{def} [\![\varphi]\!]_{M,f} = 1$

So the two notions are interchangeable, but (20) offers the advantage that it allows us to refer to the semantic value of a DRS. This will become important later on, when we introduce reference markers which have the same type of values as DRSs.

It is not difficult to extend this framework so as to include the classical possible-worlds analysis of modality. To do so, we first extend the DRS language by adding the following clause to the definition in (11):

(23)   If $\varphi$ is a DRS, then $\diamondsuit\varphi$ is a DRS-condition.

As far as accessibility is concerned, the possibility operator has the same status as the negation sign. Therefore, (12) is extended with the following clause:

(24)   If $\diamondsuit\psi \in \mathrm{Con}(\varphi)$, then $\varphi \leq \psi$

If we want to incorporate a possible-worlds semantics of this operator into our first version of the DRS semantics as given (16), it is clear how we must proceed: instead of stating what it means for a function to embed a DRS into a model, we must define what it means for a function to embed a DRS into a model at a given world. That is, instead of defining $M \vDash_f \varphi$, we must define $M \vDash_{w,f} \varphi$, for any model M, function $f$, DRS $\varphi$, and world $w$ in M. If we translate this into our second version of the DRS semantics, $[\![.]\!]$ must be defined relative to pairs of worlds and embedding functions, and assign each DRS a set of such pairs. I call such pairs 'states'. A state may be viewed as a partialized world, with the embedding function selecting a (typically small) number of individuals that inhabit this world.

The semantics for modal DRT may now be defined thus (here and in the following I suppress the subscript 'M'):

(25)     *Semantics for modal DRT*

Let $M = \langle W, D, I \rangle$ be a model, where W is a set of worlds, D is a domain of individuals, and I is an interpretation function. Let $s = \langle w, f \rangle$ be a state.

a.     $[\![\varphi]\!]_s = \{\langle w, g \rangle \mid f \subseteq g$ and $\text{dom}(g) = \text{Acc}(\varphi)$ and
       $\forall \psi \in \text{Con}(\varphi): [\![\psi]\!]_{\langle w, g \rangle} = 1\}$

b.     $[\![Pu_1 \dots u_n]\!]_s = 1$ iff $\langle f(u_1), \dots, f(u_n) \rangle \in I_w(P)$

c.     $[\![u = v]\!]_s = 1$ iff $f(u) = f(v)$

d.     $[\![\neg\varphi]\!]_s = 1$ iff $[\![\varphi]\!]_s = \varnothing$

e.     $[\![\varphi \vee \psi]\!]_s = 1$ iff $[\![\varphi]\!]_s \cup [\![\psi]\!]_s \neq \varnothing$

f.     $[\![\varphi \Rightarrow \psi]\!]_s = 1$ iff $\forall s' \in [\![\varphi]\!]_s : [\![\psi]\!]_{s'} \neq \varnothing$

g.     $[\![\Diamond\varphi]\!]_s = 1$ iff $\exists w' \in W: [\![\varphi]\!]_{s'} \neq \varnothing$, where $s' = \langle w', f \rangle$

Thus each DRS $\varphi$ determines a set of states, i.e. $\{s' \mid \exists s: s' \in [\![\varphi]\!]_s\}$, as well as a classical proposition, i.e. a set of worlds: $\{w \mid \exists s, f: \langle w, f \rangle \in [\![\varphi]\!]_s\}$.

This is the DRS language that we will work with in this chapter. In later chapters I will argue that it is deficient in certain respects, and extend it accordingly.


## 2.2 From anaphora to presupposition

Suppose we have just been told that Fred has a rabbit, and the speaker goes on to claim that 'Fred's rabbit is pink'. His utterance presupposes that Fred has a rabbit, but intuitively we would also like to say that the speaker intends the definite NP *Fred's rabbit* to refer to the *same* individual he introduced earlier on in the discourse. It is part of the traditional wisdom about definite NPs that they may be used anaphorically, to refer back to an object that is already part of the conversational context. Van der Sandt (1992) proposes to generalize this idea to the whole class of presupposition-inducing expressions, so that, in effect, the notion of presupposition comes to subsume that of anaphora, and shows that this generalization results in a very natural solution of the projection problem.

The first one to see that there is a close tie between presupposition and anaphora may have been Kripke. The penultimate footnote of Soames's 1989 paper records Kripke's observation that 'presupposition creating elements may, in some way, be anaphoric with other elements in the discourse or context.' (Soames 1989: 614) Soames gives examples with *too*, *again*, and *stop*, but it remains unclear how widely Kripke meant his observation to

apply. Soames reports that the final version of his paper was completed in 1986, which is one year before van der Sandt noted the link between presupposition projection and anaphora resolution (van der Sandt 1987). However, although it can hardly be called a detailed exposition of the theory he was going to publish some years later,[3] van der Sandt's 1987 paper already contains the gist of that theory. If Kripke was the first to see that presupposition and anaphora are allied concepts, van der Sandt was the first to realize how this insight could be turned into a theory. As far as one can tell from Soames's note, Kripke saw the connection but had no theory to account for it.

Van der Sandt's theory is based upon the observation that there are systematic parallels between anaphoric and presuppositional phenomena. Once they have been pointed out these parallels are quite striking, and it remains something of a mystery to me why they took so long to be discovered, especially when one considers that quite a few people had worked on both problems long before Kripke and van der Sandt saw that they are related. The following catalogue is to demonstrate the parallels between anaphora and presupposition, and to convey something of my wonderment at their having managed to escape notice for so long.

Let us begin with conjoined sentences. If we think of NPs like *just one man* or *many sheep* as quantified expressions in the logicians' sense, examples such as (26a, b) become problematic, and cases like these have been widely discussed in the anaphora literature. (27a-c) are parallel examples from the presupposition literature.

(26)   a.   Just one man broke the bank at Monte Carlo, and *he* has recently died a pauper. (Geach 1962)

       b.   John owns many sheep and Harry vaccinated *them* last July. (Evans 1977)

(27)   a.   James went to the circus and he *regretted* going to the circus. (Kempson 1977)

       b.   Someone has solved the problem and *it's* Lauri *who* has solved it. (Soames 1982)

       c.   Joan proved the theorem and Martha proved it, *too*. (Soames 1982)

Traditionally, the problems posed by these two sets of examples would be described in entirely different terms. In the anaphoric cases, the problem was to explain how a pronoun in the second half of a conjunction can have a quantified antecedent in the first half. In the presuppositional cases, the

---

[3]   See van der Sandt and Geurts (1991) and van der Sandt (1992).

problem was to account for the intuition that a presupposition triggered in
the second conjunct may be 'blocked' or 'suspended' if its content is
explicitly asserted in the first conjunct; a speaker who utters a sentence of the
form '$\varphi$ and $\psi\{\varphi\}$' does not presuppose $\varphi$, he asserts it. But if anaphora is a
special case of presupposition, there is really only one problem here, not
two.

In the following examples we see the same pattern recurring again and
again: on the one hand there are cases that for a long time were viewed
exclusively as instances of anaphora, while on the other hand there are
analogous examples which used to be discussed in presupposition theory.
This analogy is explained, as we will see, on the assumption that anaphora is
a species of presupposition. I write 'anaphora' and 'presupposition' in scare
quotes as a reminder that these phenomena used to be relegated to separate
departments of the semantics/pragmatics enterprise.


*Conjunctions in intensional contexts*

'Anaphora':

> John wishes to catch a fish and eat *it*. (Montague 1973)

'Presupposition':

> Maybe Mary proved the theorem and John proved it, *too*. (Soames
> 1982)
>
> It is possible that Fred has managed to kiss Cecilia and that he will
> kiss her *again*. (Karttunen 1973a)
>
> Mary thinks that she has no brothers and sisters and that her
> parents *regret* that they have only one child (McCawley 1981)


*Conditionals*

'Anaphora':

> If Smith owns a donkey, he beats *it*. (Geach 1962)

'Presupposition':

> If Haldeman is guilty, then Nixon is guilty, *too*. (Soames 1979)
>
> If Jack has children then *all* of Jack's children are bald. (Karttunen
> 1973a)
>
> If Fred has managed to kiss Cecilia, Fred will kiss Cecilia *again*.
> (Karttunen 1973a)
>
> If the problem has been solved, *it was*n't Alex *who* solved it.
> (Soames 1979)

*Disjunction*

'Anaphora':

> Either Smith does not own any donkey or he beats *it*. (Geach 1962)

'Presupposition':

> Either there is no king of France, or *the king of France* is in hiding. (Soames 1979)

> Either Jack has no children or *all* of Jack's children are bald. (Karttunen 1973a)

> Either Geraldine is not a Mormon or she has *given up* wearing her holy underwear. (Karttunen 1973a)

*Modal subordination*

'Anaphora':

> You must write a letter to your parents. *It* has to be sent by airmail. (Karttunen 1976)

> It's certain that you'll find a job, and it's conceivable that *it* will be a good-paying one. (McCawley 1981)

'Presupposition':

> It is possible that John has children and it is possible that *his children* are away. (Gazdar 1979)

> Possibly Boris killed Louis and possibly Boris *regrets* killing Louis. (Gazdar 1979)

*Hob-Nob sentences and related phenomena*

'Anaphora':

> Hob thinks a witch has blighted Bob's mare, and Nob wonders whether *she* (the same witch) killed Cob's sow. (Geach 1967)

> Tom expects to catch a fish and intends to fry *it* for dinner. (McCawley 1981)

'Presupposition':

> Bill believed that Fred had been beating his wife and Harry hoped that Fred would *stop* beating her. (Gazdar 1979)

> Bill believed that Fred had been beating his wife and hoped that Fred would *stop* beating her. (Karttunen 1973a)

It will be evident, I trust, that there are significant parallels between the anaphora and the presupposition cases. Van der Sandt's explanation for these parallels is there aren't really two phenomena to begin with: anaphora is a species of presupposition.[4] In the following pages I will outline how this hypothesis can be fleshed out so as to account for the more straightforward cases. The less straightforward ones involving modals and attitude contexts will be addressed in later chapters.

Van der Sandt's theory is an extension of DRT, and consists of three principal claims. The first of these is that anaphora is a species of presupposition, and that the standard presupposition-inducing expressions (such as definite NPs, factives, aspectual verbs, and so on) differ from pronominal anaphors mainly in that they possess a richer semantic content. This difference explains why the standard presupposition inducers, unlike anaphoric pronouns, may be interpreted by way of accommodation, which is the second key notion in the theory. Finally, van der Sandt assumes that the process of accommodation is subject to certain constraints.

Formulated in procedural terms, van der Sandt's theory predicts that if an utterance contains a presupposition-inducing element, the hearer will initially attempt to bind the presupposition to a suitable antecedent, just as he would try to bind an ordinary anaphor. If the presupposition cannot be so bound, it will be accommodated, i.e. it will be inserted in some accessible DRS. In general the number of positions at which a presupposition may be accommodated is greater than one, and van der Sandt proposes that the choice between these alternative positions is restricted by general constraints on the pragmatic acceptability of DRSs.

Before we have a closer look at this theory, I want to settle some terminological matters. I will call lexical expressions or syntactic constructions that give rise to presuppositions 'presupposition inducers' or 'presupposition triggers', and say that, for example, a definite NP 'induces' or 'triggers' a presupposition. I shall do my best to restrict my use of the term 'presupposition' to pieces of semantic information that are triggered by presupposition inducers in what I will call the 'home DRS' of the presupposition. Thus, if a definite NP occurs in the consequent of a conditional, the home DRS of the presupposition triggered by the NP will be the DRS that corresponds to the consequent of the conditional. This is the part of my terminology that I shall try to be strict about. No such attempts will be made regarding the *verb* 'to presuppose'. Following Stalnaker, I consider presupposing to be something *speakers* do, but this will not deter

---

[4] Actually, van der Sandt's (1992) claim is that presuppositions are anaphors, but this is just a terminological difference. I prefer to reserve the term 'anaphora' for referring to a subclass of presuppositions, and associate it with descriptive poverty and binding to discourse entities currently in focus. In van der Sandt's vocabulary, the term appears to have lost such connotations. See § 2.5 for further discussion.

me from saying that utterances or even sentences presuppose certain things. Whenever such locutions occur, they serve as a shorthand for more verbose formulae, like: 'In uttering this sentence a speaker would, all things being equal, presuppose such and such.'

With these terminological preliminaries out of the way, I proceed to illustrate the theory proposed by van der Sandt.

(28)   a.   If Fred has a rabbit, then Fred's rabbit is pink.

       b.   If Fred's rabbit is pink, he must be happy.

       c.   Either Fred doesn't have a rabbit or Fred's rabbit is in hiding.

Each of these sentences contains the presupposition-inducing expression *Fred's rabbit*. However, of these three (28b) is the only one from which we would infer that the speaker presupposes that Fred has a rabbit. Someone who utters (28b) presupposes that Fred has a rabbit, but if he utters (28a) or (28c) he does not commit himself to this assumption. On van der Sandt's account, these data are explained as follows. Let us assume, to begin with (28a), that at some stage in the interpretation of a token of this sentence, the following DRS has been constructed:

(29)    $[_1 : [_2$ x, y: <u>Fred x</u>, rabbit y, x owns y] $\Rightarrow$
         $[_3$ z, w: <u>Fred z</u>, <u>rabbit w, z owns w</u>, pink w]]

The underlining in (29) is to reflect that the antecedent of the conditional contains one presupposition, while the consequent contains two, one of which is embedded in the other. Going from left to right,[5] we first attempt to link the information in [x: Fred x] to an antecedent, but since no antecedent is available to this presupposition, it will have to be accommodated. There are two DRSs in which the presupposition might be accommodated: $(29_1)$ and $(29_2)$ itself, which is the presupposition's home DRS. Since in general there is a preference for accommodating presuppositions in the least embedded DRS, the presupposition will be accommodated in $(29_1)$:

(30)    $[_1$ x: Fred x,
         $[_2$ y: rabbit y, x owns y] $\Rightarrow$
         $[_3$ z, w: <u>Fred z</u>, <u>rabbit w, z owns w</u>, pink w]]

Next we try to link the information in [z: Fred z] to a suitable antecedent, which has just been set up in the main DRS. Thus we obtain (31a), which is equivalent to (31b):

---

[5]  It is natural to process presuppositions bottom-up (i.e., more deeply embedded presuppositions first) and from left to right, and this is how I will generally proceed. However, although this ordering is part of van der Sandt's original theory, it is not stipulated in the version that I present below.

(31)     a.     $[_1$ x, z: z = x, Fred x, Fred z,
                 $[_2$ y: rabbit y, x owns y] $\Rightarrow$
                 $[_3$ <u>w</u>: <u>rabbit w, z owns w</u>, pink w]]

         b.     $[_1$ x: Fred x,
                 $[_2$ y: rabbit y, x owns y] $\Rightarrow$
                 $[_3$ <u>w</u>: <u>rabbit w, x owns w</u>, pink w]]

Finally, the last remaining presupposition, [w: rabbit w, x owns w], is bound in $(31b_2)$, so that we end up with (32a), which is equivalent to (32b):

(32)     a.     [x: Fred x,
                 [y, w: w = y, rabbit y, x owns y, rabbit w, x owns w] $\Rightarrow$
                 [: pink w]]

         b.     [x: Fred x, [y: rabbit y, x owns y] $\Rightarrow$ [: pink y]]

This DRS adequately represents the preferred interpretation of (28a). In particular does it not follow from (32) that Fred has a rabbit, and thus it is correctly predicted that someone who utters (28a) does not thereby commit himself to the assumption that Fred has a rabbit. The presupposition which originates in the consequent of the conditional is bound in the antecedent, just as a personal pronoun might have been bound, and thus is 'blocked': the matrix sentence doesn't inherit the presupposition that Fred has a rabbit.

(28b) contains three presupposition triggers (for pronouns are treated as ordinary presupposition-inducing expressions), and its initial representation is the following:[6]

(33)     [: [<u>u</u>, <u>v</u>: <u>Fred u</u>, <u>rabbit v, u owns v</u>, pink v] $\Rightarrow$ [<u>w</u>: <u>male w</u>, happy w]]

Neither of the presuppositions triggered in the antecedent of (33) can be bound, and both will therefore have to be accommodated. As in the first example, there are two DRSs in which these presuppositions might be accommodated, and as before it is predicted that they will be accommodated in the principal DRS, which yields:

(34)     [u, v: Fred u, rabbit v, u owns v, [: pink v] $\Rightarrow$ [<u>w</u>: <u>male w</u>, happy w]]

The remaining presupposition is (in fact, must be) bound at top level, and thus the interpretation of (28b) results in (35a), which is equivalent to (35b):

---

[6]  In (33) the pronoun *he* is represented by the presupposition [w: male(w)], whereas in the foregoing it was assumed that the anaphor merely introduced a new reference marker. I consider the former representation to be the correct one, because pronouns generally have at least some descriptive content (but see Bosch, 1983: 151–157, for a discussion of what he calls 'pure', i.e. semantically empty, anaphors). However, I shall only use the official representation of pronouns when it helps to bring home some point or other — in this case: that anaphoric pronouns are just like other presuppositional expressions.

(35)       a.     [u, v, w: w = u, Fred u, rabbit v, u owns v, male w,
                     [: pink v] ⇒ [: happy w]]

           b.     [u, v: Fred u, rabbit v, u owns v, male u,
                     [: pink v] ⇒ [: happy u]]

This DRS *does* entail that Fred has a rabbit, and thus the theory accounts for the observation that a speaker who volunteers (28b) commits himself to the assumption that Fred has a rabbit.

If we imposed no further restrictions on the notion of accommodation we would predict that, whenever a presupposition cannot be bound to a suitable antecedent, it will be accommodated in the least embedded position, which will always be in the principal DRS. (28c) serves to illustrate that this would not always be correct. Let us start out from the following DRS for (28c), in which the presupposition induced by the proper name has already been dealt with:

(36)       [x: Fred x,
              [: ¬[y: rabbit y, x owns y]] ∨ [z: rabbit z, x owns z, in-hiding z]]

Again, there is no suitable antecedent accessible to the presupposition, which therefore will have to be accommodated. But this time, if we should allow for accommodation in the principal DRS, we would get the wrong result:

(37)       [x, z: Fred x, rabbit z, x owns z,
              [: ¬[y: rabbit y, x owns y]] ∨ [: in-hiding z]]

(37) entails that Fred owns a rabbit, which means that (28c) would presuppose (28d), and that, we observed, is not the case.

Van der Sandt proposes to deal with this problem by assuming that accommodation is subject to certain general constraints on interpretation. He defines a notion of acceptability which rules out a DRS like (37) as being pragmatically unacceptable, and stipulates that accommodation may not result in an unacceptable DRS. In the case of (36) this forces the presupposition to be accommodated in its home DRS rather than in the principal DRS, and thus it is explained why (28c) doesn't presuppose that (28d) is true.

In the following I will present an account of presupposition projection which takes as its point of departure van der Sandt's thesis that presupposition is to be analysed in terms of binding. I will accordingly call it the 'binding theory' of presupposition. I am aware that syntacticians have been using this term in a rather different sense, but their usage is so different from mine that my appropriation of the term is unlikely to engender confusion. There are three points at which my theory diverges from van der

Sandt's. First, it features a binding mechanism for presuppositions that is slightly different from the one proposed by van der Sandt. Secondly, it assumes that the constraints which, according to van der Sandt, only restrict the process of accommodation in fact restrict the binding interpretation of presuppositions as well. In this respect the binding theory is more general than van der Sandt's original proposal. Thirdly, and most importantly, the binding theory differs from van der Sandt's with respect to the nature of these constraints on presupposition projection (or, for him, accommodation).

I have shown elsewhere how the binding theory can be spelled out in a fully explicit and rigorous manner (Geurts 1995), and I will not repeat that exercise here, because I feel that in this particular case the dictum that God is in the details does not apply. As a matter of fact, the details are tedious and of little interest, and tend to deflect one's attention from more important matters. For this reason, I will abstract away from some of the technicalities that a more rigorous presentation would have to include. It is inevitable that this will lead to some loss in precision, but I hope that this loss is redressed by a gain in clarity.

The first thing to do is decide is how presuppositions are going to be represented. Thus far I have marked presuppositions by underlining parts of DRSs, and it is natural to assume that presuppositions are formally represented by DRSs; in fact, I have already adopted this representation in the foregoing. For example, the presupposition triggered in the second disjunct in (36) would be [z: rabbit z, x has z]. There are two things to note about this representation. First, as a rule the DRSs that represent presuppositions will be not be proper DRSs, because they contain reference markers that are free. The reason for this, of course, is that presupposition and assertion are *aspects* of the content of an expression which are ususaly linked and cannot be neatly separated. Secondly, and more importantly, if presuppositions are represented by DRSs it follows that their content is merely existential, or descriptive; the semantic content of [z: rabbit z, x has z], as determined by the semantics of the DRS language, is that there is an $a$ such that $a$ is a rabbit and $b$ owns $a$, where $b$ is whatever individual x has been assigned to. However, this should not be taken to imply that, according to the binding theory, there are only existential presuppositions. On the contrary, the theory claims that presuppositions are never merely existential; by definition, a presupposition wants to be bound to an antecedent. It is just that this 'definiteness' is not explicitly encoded in the content of a presupposition; it is implied by the way presuppositions are handled by the theory's construction rules.

Having decided that a presupposition will be represented by a DRS, we must settle how it is to be linked to its home DRS, i.e. the DRS in which it is triggered. The graphical representation that I have used (and will continue to use) suggests that a presupposition is *part* of its home DRS. If we take this

line, presupposition projection will have to be implemented by means of transformations on DRSs.[7]

(38)    a.    Ada will not eat mud again.

          b.    $[_1$ x: Ada x, $\neg[_2$ : x will eat mud, <u>x has been eating mud</u>]]

          c.    $[_1$ x: x has been eating mud, Ada x, $\neg[_2$ : x will eat mud]]

In (38a), the particle *again* triggers the presupposition that Ada has been eating mud (before some given date), which will be accommodated in the principal DRS, as shown in (38c). If initially this presupposition is contained in its home DRS, accommodation requires that we delete it from its home DRS and insert it in the main DRS. This is the procedure that was adopted in earlier implementations of the theory (van der Sandt 1992, Geurts 1995). However, there is an alternative procedure, which at least from a technical point of view is more elegant. We can use so-called 'underspecified' DRSs to represent the connection between a presupposition and the place where it is triggered.[8] On this account, instead of saying that the presupposition [: x has been eating mud] is contained in $(38b_2)$, we merely say that it is contained in a DRS that is accessible to $(38b_2)$, which in this case comes down to saying that it is part of either $(38b_1)$ or $(38b_2)$. The advantage will be clear: in order to get from (38b) to (38c) we merely have to *add* that the presupposition is contained in $(38b_1)$. We don't have to delete parts of the initial DRS. In general, if we use underspecified DRS to represent presuppositions, we can implement projection as a strictly incremental process, which only adds information.

It may be that, from a technical point of view, these two ways of linking presuppositions to their home DRSs are about equally complex; for although the underspecification view makes for a neater formulation of the construction component, maintaining underspecified DRSs is not as simple as one might hope. Nonetheless I tend to favour the underspecification view for conceptual reasons. If we adopt the transformational approach, presuppositions are thought of as entities that arise in one site and thence migrate to another. If we adopt the underspecification approach, on the other hand, presuppositions don't move in any sense, and it seems to me that this picture is to be preferred, if only because it seems to be more realistic from an intuitive point of view.

In the following I will nonetheless continue to pretend as if presuppositions are moved about, just because I find this a convenient manner of speaking. But I should like to stress that this is a manner of speaking only, because the

---

[7]   Beaver (1997) apparently assumes that this transformational approach is inherent to the binding theory. This is not the case, as we will presently see.

[8]   See Reyle (1993) for an underspecified version of DRT.

foregoing considerations show that the binding theory does not require that presuppositions be viewed as movable objects.

The next task on the agenda is to define what it means for a presupposition to be bound or accommodated. Following van der Sandt, I prefer to take a somewhat narrow-minded view on these notions, and keep them separate from considerations having to do with coherence, plausibility, and so on. I will later present some arguments in support of this view, but for the time being it is important to keep in mind that I use the terms 'binding' and, in particular, 'accommodation' in a more austere sense than some other people have done.

If a presupposition $\chi$ is bound, then the reference markers in $U(\chi)$ are linked to reference markers in $Acc(\varphi)$, where $\varphi$ is $\chi$'s home DRS. This linking is encoded by a set of equations $\{u_1 = v_1, ..., u_n = v_n\}$, where $u_1, ..., u_n \in U(\chi)$, and $v_1, ..., v_n \in Acc(\varphi)$. Accommodation may be viewed as a special case of binding, when the set of equations is empty.

(39)  *Binding sets*

Let $\chi$ be a presupposition whose home DRS is $\varphi$. Then a binding set for $\chi$ is a set $\{u_1 = v_1, ..., u_n = v_n\}$, where $u_1, ..., u_n \in U(\varphi)$, and $v_1, ..., v_n \in Acc(\psi)$

As is usual, it is assumed here that binding sets are simply given. Being a version of DRT, the binding theory imposes restrictions on what bindings are possible, but it does not make predictions about how hearers decide between alternative antecedents.

(40)  *Binding*

Let $\chi$ be a presupposition whose home DRS is $\varphi$, $\varphi' \leq \varphi$, B a non-empty binding set for $\chi$, and $\chi' = \langle U(\chi), Con(\chi) \cup B \rangle$. Then $\chi$ is bound in $\varphi'$ by merging $\chi'$ and $\varphi'$.

(41)  *Accommodation*

Let $\chi$ be a presupposition whose home DRS is $\varphi$, and $\varphi' \leq \varphi$. Then $\chi$ is accommodated in $\varphi'$ by merging $\chi$ and $\varphi'$.

(42)  *Projection*

Let $\chi$ be a presupposition whose home DRS is $\varphi$, and $\varphi' \leq \varphi$. Then $\chi$ is projected to $\varphi'$ if it is either bound or accommodated in $\varphi'$.

For convenience, I assume that either all or none of the reference markers in $U(\chi)$ are bound; in the former case we have binding, in the latter accommodation. This is an innocuous simplification because in the great majority of cases, a presupposition's universe contains just a single reference marker.

We are now all set to formulate the three principles that constitute the binding theory of presupposition:

(A) Presuppositions must be projected.

(B) Binding is preferred to accommodation.

(C) A presupposition must be projected to the highest possible DRS.

It will be evident that none of these principles is absolute, although the first two may be more absolute than the third one. They are all subject to general constraints on interpretation, which require that an interpretation should be consistent, coherent, and so on. These constraints and their effects on presupposition projection will be discussed in the next section.

Before these principles come into play, presuppositions are merely representational structures, and are therefore completely inert. Principle (A) drives away this inertia by requiring that presuppositions must either be bound or accommodated. Principle (B) captures the insight, discussed in § 1.2, that accommodation is a repair strategy. In principle, a presupposition wants to be bound, but if it cannot be bound it will be accommodated. Principle (C) may be viewed as a generalization of a constraint first proposed by Heim (1983). Heim distinguishes between two types of accomodation: global and local. In terms of the present framework, a presupposition is accommodated globally if it is accommodated in the principal DRS, and locally if it is accommodated in its home DRS. Heim's proposal is that, in general, global accommodation is preferred to local accommodation, and principle (C) generalizes this in two ways. First, this principle applies not only to accommodation but to projection in general. As we will see, this makes some difference from an observational point of view (though not much), and it is surely more attractive conceptually speaking. Secondly, although it is possible to capture Heim's distinction between global and local accommodation in our framework, the distinction as such doesn't play a role in the theory. In general, there is a chain of accessible DRSs in which a presupposition can be accommodated, and the two ends of this chain are the main DRS and the presupposition's home DRS. Global and local accommodation are just convenient labels for referring to accommodation in these DRSs, but they do not denote special processes.[9]

Although it is widely agreed that principle (C), or something like it, must be assumed, it is not obvious what the rationale underlying this principle might be. Note, first, that projection to the main DRS will often, though not

---

[9]  It has been suggested by Beaver (1993) and Chierchia (1995) that there are methodological reasons for preferring theories that only use global accommodation. This may be true for the type of theory that Beaver and Chierchia envisage, but in the present framework this restriction would be clearly ad hoc, and is therefore to be avoided on methodological grounds. Apparently, what is good for one type of theory need not be good for another.

always, yield stronger readings than accommodation in an embedding position. I find it doubtful, however, that hearers generally try to make their interpretations as strong as possible, and therefore I don't see that principle (C) can be motivated solely on this basis. It seems to me that this preference should be explained in terms of relevance rather than quantity (to adopt the Gricean jargon for the occasion). Consider the following example:

(43)     a.    Fred didn't bring his wife.

         b.    [x, y: Fred x, y is x's wife, ¬[: x brought y]]

         c.    [x, y: Fred x, ¬[y: y is x's wife, x brought y]]

(43a) may well be uttered in a context where it hasn't been established yet that Fred has a wife. A hearer who isn't informed about Fred's marital status will typically infer from an utterance of (43a) that Fred is married, i.e. he will accommodate Fred's wife globally, as in (43b), rather than locally, as in (43c). Why? Presumably because on the latter interpretation (43a) would generally be a pointless thing to say. To be sure, one can imagine situations in which the message represented by (43c) would be of interest, but such situations would appear to be rather marginal.

Yet another possible way of motivating principle (C) is that it will often yield readings that are intuitively simpler than possible alternatives. In some, presumably 'illogical', sense, (43c) not only is less informative than (43b); it is also more complex. If this is correct, it may help to explain why principle (C) holds.

I have briefly indicated three ways of explaining why hearers should prefer to project presuppositions to the highest possible DRS. These explanations do not exclude one another, and it could turn out that some combination of them is correct. This is as it may be. I don't have a more definite proposal to make; I merely wanted to suggest some possible reasons why principle (C) is a plausible one. That it is a plausible principle is not at issue; for, as I noted already, there can hardly be any doubt that something along these lines is required.

### 2.3  Constraints on interpretation

As a rule, people are remarkably tolerant when it comes to interpreting each other's verbal behaviour. When one interprets an utterance, he assumes that what the speaker wants to say makes sense and is useful, given the purposes of the conversation. This mundane observation is the starting point of Grice's seminal work on pragmatic reasoning. As is well known, Grice presents his theory as a collection of conversational maxims, which he takes to be instances of a single super maxim, viz. that speakers should be cooperative. From the hearer's point of view, these maxims act as constraints on the

interpretation of the speaker's contribution. As long as he can, the hearer will assume that speaker is trying to be cooperative, and this assumption constrains his interpretation strategy in various ways.

In (44) below I propose three general constraints on interpretation, two of which are entailed by Grice's theory of conversation; the third constraint does not correlate with anything in the Gricean framework, since it is a purely formal restriction on the class of DRSs that count as acceptable interpretations. This set of constraints is not meant to be complete in any sense. In this section I am interested in the way general constraints on interpretation interact with presupposition projection, and for this purpose the set in (44) will be sufficient. It should be clear, however, that it is not sufficient for all purposes, nor do I want to exclude the possibility that there are projection phenomena that interact with constraints not listed here.

In order to interpret a discourse the hearer maintains a DRS which he updates as the speaker's delivers his utterances. The speaker utters a sentence $S_i$, and the hearer extends and, perhaps, revises his current DRS $\varphi_i$, thus arriving at a new DRS $\varphi_{i+1}$. This process of extension and revision is subject to at least the following restrictions:

(44)    *General constraints on interpretation*

    a.    *Consistency*: $\varphi_{i+1}$ must be consistent.
            Cf. Grice's second quality maxim: 'Do not say what you believe to be false.'

    b.    *Informativeness*: $\varphi_{i+1}$ must be informative, i.e. $\varphi_{i+1}$ may not be entailed by $\varphi_i$.
            Cf. Grice's first quantity maxim: 'Make your contribution as informative as is required (for the current purposes of the exchange).'

    c.    *Well-formedness*: $\varphi_{i+1}$ must be a proper DRS.

I take it that there is no need to justify these constraints, so I will only make a few explanatory remarks on the Gricean constraints, i.e. (44a) and (44b). First, these constraints are defeasible. As Grice would put it, in some situations conversational rules can be exploited by breaking them. Thus a speaker may make a point by saying something contradictory, for example:

(45)    A:    Mr. Schwarzenegger should run for president.

        B:    Yeah, and four is a prime number.

Here B expresses his disagreement by uttering a contradiction. Similarly, the informativeness constraint is violated, for example, whenever a speaker repeats a statement to lend it emphasis.

These constraints act upon the interpretation process through the hearer's presumption that they will not be violated. Accordingly, the hearer tries to come up with an interpretation which violates as few of these constraints as possible. However, the Gricean constraints will sometimes pull in opposite directions. It may happen that the hearer must choose between an interpretation which violates the consistency constraint and another reading which violates the informativeness constraint. I will assume that, in such a situation, the consistency constraint takes priority over the informativeness constraint. This accords with the familiar observation that the quality maxim has a special status among the maxims: it is more urgent for a statement to be consistent than to be informative (cf. Grice 1981, Gazdar 1979: 46–48).

There are many respects in which the constraints in (44) regulate the interpretation of an utterance, and presupposition projection is just one of many processes affected. The principles of the binding theory interact with these constraints in different ways, some of which are more interesting than others. Let me begin with a less interesting case. In the following examples, there are conflicts between principle (A), which says that presuppositions must be project, and the consistency and informativeness constraints, respectively:

(46)    a.    ?France is a republic. The King of France is bald.

        b.    ?The King of France is bald. The King of France is bald.

In (46a) the presupposition cannot be bound, obviously, and it cannot be accommodated without violating the consistency constraint. Similarly, the presupposition triggered by the second occurrence of *the King of France* in (46b) can neither be bound nor accommodated without violating the informativeness constraint. Intuitively, I should say that in both cases projection 'succeeds', and Gricean constraints are consequently violated, but I set no great store by intuitions about this sort of example (cf. § 1.6), and henceforth I will assume for convenience that principle (A) of the binding theory is not defeasible.

Things become more interesting when we consider principle (B), which says that binding is preferred to accommodation. The following example shows this principle to be defeasible by general constraints on interpretation:

(47)    Fred's wife will be at the reception, so Barney will be sorry that his wife can't come.

Suppose the hearer doesn't know that Barney is married. In principle, the presupposition triggered by *his wife* may therefore be bound to the reference marker representing *Fred's* wife. But if the presupposition were so bound, the resulting interpretation would be inconsistent, and therefore an

accommodation reading is preferred in this case. So, principle (B) may have to give in to save consistency.

Things really become interesting when we turn to principle (C) of the binding theory, and in the remainder of this section I will discuss some of the ways in which this principle interacts with Gricean constraints on interpretation (the interplay between principle (C) and the well-formedness constraint will loom large in §§ 2.4.1 and 2.4.4). The workings of principle (C) are best viewed in procedural terms. According to this principle, the hearer will always first try to project a presupposition to the main DRS, but if the resulting interpretation violates an interpretation constraint, he goes down the path of accessible DRSs until a DRS is found which will accept the presupposition without violating any constraint.

(48)  a.  Harry is a bachelor, so it wasn't Harry's wife who shot the burglar.

  b.  [x: Harry x, bachelor x, ¬[u: u is x's wife, u shot the burglar]]

  c.  [x, u: Harry x, bachelor x, u is x's wife, ¬[: u shot the burglar]]

Principle (C) says that the underlined presupposition in (48b) is preferably accommodated in the principal DRS, which would yield (48c). However, on this interpretation (48a) would be contradictory, so on the first run principle (C) is defeated. Given that accommodation in the main DRS is not possible, the principle says that accommodation one level down is the preferred option, and in this case it is the only remaining possibility, too, so the reading we predict is (48b), which is correct. Note that, on this reading, (48a) still violates the informativeness constraint, because the second half of the sentence is entailed by the first half. If (48b) is nonetheless the preferred reading of (48a), it is because consistency takes precedence over informativeness.

In (48a), the presupposition that Harry has a wife is blocked because it is inconsistent with the speaker's explicit claim that Harry is a bachelor. This is an exceptionally blatant sort of inconsistency. A rather more subtle type of inconsistency threatens when a presupposition contradicts, or would contradict, what the speaker has said or implied *about* his own knowledge. (49) is a case in point:

(49)  a.  I don't know if Harry is married ...

  b.  but the woman we saw him with last night certainly wasn't his wife.

By uttering (49) the speaker does not presuppose that Harry has a wife, although (49b) contains a definite NP that triggers this presupposition. Hence the presupposition that Harry has a wife is prevented from accommodating in the principal DRS, and is forced to accommodate within

the scope of the negation instead. Why? The answer is fairly obvious: if the presupposition had been accommodated at top level, the resulting DRS would contradict the speaker's explicit claim that he doesn't know if (50) is true.

(50)      Harry is married.

Let us introduce the word *issue*, as a term of art, to refer to propositions expressible by sentences of the form '*a* does not know if φ is true'. Such an issue is *settled* as soon as the same speaker commits himself to the truth or falsity of φ. Thus somebody who utters (49a) introduces into the context an issue in the most direct and explicit way possible, but issues may also be conveyed in a more implicit manner.

(51)      a.    It is possible that Harry is married ...

          b.    but the woman we saw him with last night certainly wasn't his wife.

In (51a) the speaker doesn't *say* in so many words that he doesn't know if (50) is true, but nonetheless we infer from his utterance that this is so. This type of inference is standardly construed as a conversational implicature, and this diagnosis is corroborated by the observation that it is cancellable:

(52)      a.    It is possible that Harry is married ...

          b.    In fact, I'm pretty sure that he is married.

But although this inference is cancellable, it is not cancelled by the presupposition triggered by the definite NP *his wife* in (51b). Instead, what happens is that the implicated issue remains in force and prevents the presupposition from floating up to the principal DRS, just as the asserted issue did in (49b).

My suggestion is that in (51b) local accommodation is enforced by the implicature that the speaker doesn't know if Harry is married. There is an alternative way of viewing this case, for the hearer may reason as follows: If the presupposition that Harry has a wife is accommodated in the main DRS, then in retrospect it becomes something of a mystery why he should have uttered (51a). Or, in other words, why should the speaker consider the possibility that Harry is married, only to reject it immediately afterwards? This is how van der Sandt would explain why the presupposition is accommodated locally. I have no principled objections against this view, because it seems to me that it just a different way of spelling out the very same intuition that informs my own. Nevertheless, I believe that my proposal is the more plausible one. On my account, a hearer who interprets (51a) will take this to convey that the speaker doesn't know if Harry is married, and this implicature affects the interpretation of the definite NP in the next

sentence (i.e. the implicature causes the presupposition to be accommodated locally rather than globally). On van der Sandt's account, the hearer rules out the possibility of global accommodation because it would render the speaker's previous utterance infelicitous. The main reason for preferring my account is that its *procedure* is more plausible. In my opinion, it is more likely that the construal of *his wife* in (51b) is affected by what the hearer has inferred already than that he should go back to consider (51a) in the light of a possible construal of *his wife* in (51b).

I believe that the account that I have outlined is intuitively plausible, but nonetheless one may wonder if it is really cogent. If the interpretation of (51b) involves a clash between a presupposition and an implicature, why is it the latter that wins out, not the former? It is at least conceivable that the implicature that the speaker doesn't know if Harry is married is cancelled in order to make it possible for the presupposition to be accommodated in the principal DRS. Why doesn't this happen?

Before I answer this question, let us first try to get a clearer view on what the problem is.[10]

(53)    a.    ?Maybe Fred has been drinking, but Fred has stopped drinking.

           b.    Fred has stopped drinking

           c.    Fred has been drinking.

In (53a) the presupposition that (53c) is true cannot be blocked at an embedded level of representation because it is triggered in the principal DRS to begin with. In principle, this presupposition can be accommodated, for it seems possible to utter (53b) in a context in which it is not yet part of the common ground that (53c) is true. It might be expected therefore that the implicature that the first clause of (53a) gives rise to — that the speaker doesn't know if (53c) is true — would simply be cancelled, so as to make room for the presupposition. Apparently, this doesn't happen. We perceive a distinct tension between the two conjuncts of (53a), as a result of which any utterance of this sentence will be infelicitous.

To see why this is so, consider (52) again, in which the implicature that (52a) gives rise to is cancelled *explicitly*, without however resulting in an awkward discourse. There is a clear difference between (52) on the one hand and (53a) on the other, and I believe that this difference is due to the fact that implicatures cannot be cancelled unceremoniously. Whenever a speaker wants to get rid of a conversational implicature, he should motivate or at least clearly signal his move. In (52a) and the first clause of (53a) the speaker

---

[10] The first conjunct of (53a) can perhaps be read as admitting that Theo has been drinking, but it is not this concessive interpretation that is intended here.

goes out of his way to avoid putting forward a stronger and at the same time simpler utterance, i.e. he ostentatiously fails to make the less verbose and more informative statements in (50) and (53c), respectively — and that is precisely what conveys, by way of implicature, that he doesn't know whether these simpler sentences are true. It appears, then, that a implicature that arises, inter alia, because the utterance that induces it is comparatively verbose cannot be cancelled without further ado. It requires some visible effort to get rid of such an implicature — or to it put otherwise, it requires emphasis. Note for instance that the expression *in fact* in (52b) is not redundant. It serves to signal that the speaker wants to retract an implication that his previous utterance has given rise to, and without such a signal the sequence (52a, b) would be awkward: the unadorned assertion that Harry is married is not sufficient to cancel the implicature that (52a) gives rise to. But if simple assertions don't suffice to cancel the kind of implicature that arises in (52a) and (53a), it is only to be expected that *presuppositional* devices will not be sufficient either; for presupposition surely is a less emphatic way of presenting information than assertion. Hence the infelicity of (53a).

To sum up, implicatures cannot be cancelled by presuppositional means because it requires emphasis to cancel such implicatures, and the effect of a presuppositional expression is precisely to de-emphasize information. By the same token, if accommodation of a given presupposition would be inconsistent with an implicature, the presupposition will, if at all possible, be accommodated in a non-entailed position, so that the conflict is avoided.

Thus far we have seen that the interpretation of presupposition-inducing expressions is constrained by propositions of the form '*a* does not know whether $\varphi$ is true', and I have argued that it really doesn't matter how such issues enter the context: whether they are explicitly asserted or merely implicated, their effect on the projection process is the same in either case. It should be noted that the relevant issues in this connection are the *speaker's*. What is an open issue for one person may be settled for another, and if A has introduced an issue into the context, B is not restricted by it in his utterances in the same way A is. Indeed, B may settle an issue brought up by A, simply by presupposing that it is resolved. To illustrate, there is nothing wrong with the following discourse, provided A and B are different speakers:

(54)     A:   Maybe Fred has been drinking.
         B:   Fred has stopped drinking.

Of course, this observation does not entail that issues don't constrain presupposition projection, but it does call for qualification of what I said earlier. Apparently, conflicts between issues and presuppositions can be tolerated as long as they belong with different speakers; it is only that a speaker may not presuppose $\varphi$ (or not-$\varphi$) if he himself has signalled that he doesn't know if $\varphi$ is true.

In the preceding section I used examples like (55) to show that presuppositions cannot always be accommodated globally:

(55)     Either Fred doesn't have a rabbit or Fred's rabbit is in hiding.
         (= (28c))

An utterance of (55) does not presuppose that Fred has a rabbit. This may now be explained as follows. A speaker who utters a disjunction of the form 'φ or ψ' conversationally implicates that he doesn't know the truth value of either φ or ψ. But the issue whether Fred has a rabbit would be settled if we allowed the presupposition that Fred has a rabbit to project to the principal DRS. Therefore, the presupposition must be accommodated in some other DRS, which can only be its home DRS. So the reading we predict for (55) is (56):

(56)     [x: Fred x,
             [: ¬[y: rabbit y, x owns y]] ∨
             [z: rabbit z, x owns z, in-hiding z]]

This is not yet the complete story about presupposition projection in disjunctions. I will return to this topic in § 2.4.3.

This line of explanation is based upon an idea that also underlies the theories of Gazdar (1979) and Soames (1979, 1982),[11] but the actual implementation is quite unlike either, and I want to briefly comment on what I take to be the most important differences. Although I shall restrict my remarks to Gazdar's theory, they also apply to the proposal advanced by Soames. The main point I want to make is that within the kind of framework envisaged by Gazdar, no sense can be made of the interaction between presuppositions and conversational implicatures. In Gazdar's view a presupposition is a piece of information that is discarded as soon as it becomes clear that it is incompatible with a conversational implicature. For Gazdar, the definite NP *Fred's rabbit* in (55) induces the presupposition that the speaker knows that Fred has a rabbit, but this presupposition is thrown away because by uttering this sentence the speaker conversationally implicates that he doesn't know if Fred has a rabbit. Therefore, Gazdar assumes that conversational implicatures may quite literally *cancel* presuppositions.

The problem with this type of account is that it is based upon the idea that, in examples like (55), the process of presupposition projection is driven by conflicts between implicated and presupposed information and that, whenever such conflicts arise, it is the implicatures that win the day. First, it strikes me as curious that, under otherwise ordinary circumstances, a speaker

---

[11] The Gazdar/Soames theory is foreshadowed in Stalnaker's (1974: 207–209) remarks on factive verbs.

should want to provide his audience with conflicting bits of information, trusting that they will be able to figure out for themselves which ones he is prepared to commit himself to. Secondly, why is it that conversational implicatures may cancel presuppositions and not the other way round? This question, which Gazdar doesn't answer, is especially urgent because one of the characteristic features of conversational implicatures is precisely that they are cancellable. Thirdly, it would seem that presuppositions at least logically precede conversational implicatures, because they contribute to the proposition that an utterance expresses, on a given occasion, whereas conversational implicatures are derived from such propositions. In more authentically Gricean terms: while conversational implicatures are determined *on the basis of* 'what is said', presuppositions *contribute* to what is said by a speaker on a given occasion. Of course, Grice himself would disagree with the second half of this claim, for he is evidently unhappy with the notion of presupposition, and leans towards the view that all alleged presuppositions are to be explained in different terms. It should be clear that I am not prepared to follow Grice in this — nor is Gazdar, by the way. As discussed in § 1.2, I take the common-sense view that presupposed information is (presented as) given, and if we want to accommodate this notion within a Gricean framework it is only natural to assume that it would be part of what is said.

There is a further point that should be mentioned in this connection. Grice holds that, in order to determine what a speaker says by uttering a sentence φ, any anaphors that occur in φ must be resolved first, and since I have argued that presupposition *subsumes* anaphora, the most natural place for presuppositions, again, would seem to be before 'what is said', and thus way before any conversational implicatures can arise. However, Gazdar's cancellation account requires that this scheme be turned upside down. In order to determine the full content of an utterance, according to Gazdar, its conversational implicatures must be determined first, and only those presuppositions will survive which are compatible with the utterance's literal meaning and its implicatures. As Beaver puts it:

> [Gazdar's] account does not bear any relation to the fairly intuitive notion of presuppositions as previous assumptions. Indeed, since presuppositions are the last things to be added in Gazdar's definition of update, perhaps it would be more natural to call them *post-suppositions*. (Beaver 1997: 965)

As I have tried to argue in the foregoing, the binding theory presents a more plausible picture of the interaction between presuppositions and conversational implicatures. One important factor that contributes to its plausibility is that we can maintain that presuppositions are not cancellable. Thus, a fortiori, a conversational implicature may never cancel a

presupposition, but it only may prevent presuppositional information from ascending beyond a certain level in the discourse representation.

Our general picture of the interaction between presupposition and implicature amounts to the following. When a speaker utters a sentence $\varphi$ in a context c, and $\varphi$ contains some presupposition-inducing expression $\chi$, then there will in general be several possible ways of interpreting $\chi$: the presupposition it triggers may be bound or accommodated at various positions in the discourse representation. Each of these possibilities results in a possible interpretation of $\varphi$, and each of these possible interpretations will license, or help to license, a number of conversational implicatures. On the other hand, the implicatures that a possible interpretation gives rise to may influence the hearer's judgment as to its adequacy to the speaker's intentions — whether or not it is a plausible interpretation of his utterance. It follows from this that the content of an implicature is partly determined by the interpretation of presupposition-inducing expressions. The converse, on the other hand, does not hold, although it is true that alternative interpretations of an utterance (and in particular, alternative ways of projecting presuppositions that its initial representation contains) are evaluated, inter alia, on the basis of the implicatures that they would give rise to.

In my opinion this is a plausible picture, and it certainly is a lot more plausible than the one proposed by Gazdar. Although Gazdar's idea about the interaction between presupposition and implicature was fundamentally correct, the same cannot be said of the framework in which it was implemented. Within the framework of the binding theory Gazdar's idea can be implemented in a far more natural way.

## 2.4 Applications and refinements

In this section I present a handful of small case studies with a view to bring out certain aspects of the binding theory, and to introduce some refinements.

### 2.4.1 Pronouns

Binding a presupposition $\chi$ is done in three steps. We first determine a binding set for $\chi$, then merge this set with $Con(\chi)$, and finally put $\chi$ into an accessible DRS. This entails that a presupposition need not be inserted in the same DRS in which it is bound. More accurately: a reference marker in $U(\chi)$ may be linked to a reference marker which is introduced in a higher DRS than the one to which $\chi$ is projected. This arrangement is motivated by

two considerations. First, it makes for a more elegant formulation of the binding theory than other procedures that I can think of. In particular, it makes it possible to maintain that principle (C), which says that a presupposition must be projected to the highest possible DRS, applies regardless whether the presupposition is bound or accommodated. Secondly, there are observations which suggest that something like this division of labour is called for. Suppose Fred has just introduced his friend Barney to a person named 'Leslie', who is called away immediately after the introduction (note that the English name 'Leslie' is not gender specific). Then the following conversation ensues:

(57)     Barney:     What a nice guy.

           Fred:     Whaddayamean, 'a nice guy'? Leslie is a woman!

           Barney:     If Leslie is a woman, she sure has developed her masculine side.

Barney's last utterance may be read as a concession, in which case he accepts Fred's assertion that Leslie is a woman, but it may also be read as a genuinely hypothetical claim, in which case Fred has not (yet) managed to convince his friend. It is the second construal that is relevant here. We start with the following representation:

(58)     [x: Leslie x,
               [: woman x] $\Rightarrow$
               [u: female u, u developed her masculine side]]

The presupposition [u: female u] obviously must be bound to x in the principal DRS, so its binding set is {u = x}. This set is merged with the presupposition, which yields [u: u = x, female x], and this enlarged presupposition must be projected in accordance with principle (C) of the binding theory. So we must first try to insert it in the main DRS. This is not possible, however, because Barney's utterance implicates that he doesn't know if Leslie is a woman, and therefore the presupposition lands in the antecedent of the condtional. Hence the interpretation we end up with is (59a), which, supposing that 'woman' entails 'female', is equivalent with (59b).

(59)     a.     [x: Leslie x,
                     [u: u = x, female u, woman x] $\Rightarrow$
                     [: u developed her masculine side]]

           b.     [x: Leslie x, [: woman x] $\Rightarrow$ [: x developed her masculine side]]

This is the reading that we wanted to account for. Note that it could only be derived because of the way we have factorized the binding procedure.

Our procedure for binding presuppositions deviates from the procedure adopted in standard DRT for the interpretation of pronouns, where the information contributed by a pronoun is not moved at all. According to the standard version of DRT, the conditional statement in (57) would be construed as follows:

(60)    [x: Leslie x,
            [: woman x] ⇒
            [u: u = x, female u, u developed her masculine side]]

I have not followed the standard theory in this point for two reasons. First, if the *in situ* binding of pronouns is extended to all presupposition triggers, binding and accommodation become quite different things, which is not only intuitively incorrect but would also require a more cumbersome formulation of the binding principles. Secondly, *in situ* binding of pronouns (and other presupposition triggers) leads to incorrect predictions in intensional contexts like (61a).

(61)    a.    If Barney had been a woman, Fred would have married him.

        b.    [x, y: Barney x, Fred y,
                  [: woman x] > [u: u = x, male u, y marries u]]

        b.′   [x, y: Barney x, Fred y, [: woman x] > [: male x, y marries x]]

        c.    [x, y, u: u = x, male u, Barney x, Fred y,
                  [: woman x] > [: y marries u]]

        c.′   [x, y: male x, Barney x, Fred y, [: woman x] > [: y marries x]]

Here '>' symbolizes the counterfactual conditional. If the pronoun in (61a) is bound *in situ*, according to the standard DRT procedure, we obtain (61b), which is equivalent to (61b′). This says that, if Barney had been a woman, then he would have been male and Fred would have married him, which is incorrect. If, on the other hand, we follow the rules of the binding theory, we obtain (61c), which is equivalent to (61c′), and gives the right interpretation.

In all fairness it should be noted that this argument hinges upon the assumption that the descriptive content of (English) pronouns has the same status as the descriptive content of a full definite NP, say. Since Kamp and Reyle do not make this assumption,[12] and they don't deal with counterfactuals in their book, either, (61) is strictly speaking not a

---

[12] Although Kamp and Reyle (1993: 70–73) discuss the matter at some length, they don't have an official proposal for dealing with gender, but from what they say it is clear that they prefer to view gender in grammatical terms. It seems to me that, for English pronouns at least, this is incorrect, and that the treatment suggested here is more adequate, but an in-depth discussion of this issue is not on my agenda.

counterexample against their theory. However, the purpose of this example was not so much to prove that standard DRT is wrong as to explain why the binding mechanism that I propose deviates from the standard account.

### 2.4.2 Denials

In § 2.3, it was shown how the binding theory deals with examples like the following:

(62)    a.    Harry is a bachelor, so it wasn't Harry's wife who shot the burglar. (= (48a))

        b.    It wasn't Harry's wife who shot the burglar: Harry is a bachelor.

In (62a) the presupposition triggered by *Harry's wife*. This presupposition is accommodated in its home DRS, because if it would be accommodated in the main DRS the resulting reading would violate the consistency constraint. Formulated in Heim's terminology, the presupposition is construed by way of local accommodation. Essentially the same analysis applies to (62b). The only difference between (62a) and (62b) is the order in which they present their information. As a consequence, (62a) violates the informativeness constraint, as we have seen, while (62b) does not.

Statements like (62b) are sometimes called 'denials'. Denials present problems to many theories of presupposition, and occasionally it is suggested that they are fundamentally different from ordinary (read: less problematic) instances of presupposition projection, and that, accordingly, their treatment is not so urgent. In the more recent literature, this suggestion tends to be accompanied by a further suggestion to the effect that Horn's (1989) work on 'metalinguistic negation' confirms that presupposition theorists need not worry about denials.[13] According to Horn, (62b) is just an instance of a much more general phenomenon, which has nothing to do with presupposition projection as such. The following examples (from Horn 1989: 371) are all instances of metalinguistic negation alongside cases of presuppositional denial like (62b) :

(63)    a.    He didn't call the POlice, he called the poLICE.

        b.    I didn't manage to trap two monGEESE — I managed to trap two monGOOSES.

---

[13] For extensive discussion of denials, see Horn (1989a), van der Sandt (1991), Carston (1994), and Geurts (1998a). The suggestion that metalinguistic negation offers solace to the presupposition theorist troubled by denials has been made, apart from Horn himself, by Beaver (1992), Krifka (1994), Chierchia (1995), and, interestingly, van der Sandt (1991).

c.  (Esker too ah coo-pay luh vee-and?)
    Non, je n'ai pas 'coo-pay luh vee-and' — j'ai coupé la viande.

d.  We didn't make love— we fucked.

e.  We didn't fuck — we made love.

If it is true, as Horn claims, that these and (62b) are all instances of the same phenomenon, then presuppositional denials are not a problem for a theory of presupposition projection; they can be relegated to a theory, such as Horn's, which treats them on a par with the examples in (63).

In Geurts (1998a) I have discussed and criticized Horn's ideas at great length, and here I will confine myself to a few general remarks and some specific observations about presupposition denials. My main objection against Horn's notion of metalinguistic negation is that it covers disparate phenomena; it is not a coherent concept. In particular, presupposition denials have nothing to do with the facts observed in (63). The only thing that all these cases have in common is that they are intonationally marked, and that they would typically be used to correct an utterance made by the previous speaker. But the similarity doesn't go deeper than that. To see this, note that the examples in (63) would normally be executed with a contrastive intonation contour. In (63a), for instance, 'POlice' is replaced with 'poLICE', and this is reflected in the constrastive accents that these words will get. On the other hand, (62b) isn't contrastive in any way, although normally speaking this sentence will be intonationally marked, too, for instance as follows:

(64)    It WASN'T Harry's wife who shot the burglar: Harry is a BACHELOR.

But this is not a contrastive intonation pattern.

Another way of seeing that the notion of metalinguistic negation is ill-conceived is by asking ourselves whether it will cover a case like (62a). Unlike (62b), this example can be realized without a heavily marked intonation contour, so if someone should want to claim that this is an instance of metalinguistic negation, he would have to explain why the markedness that is allegedly one of the hallmarks of this phenomenon is absent here. If, on the other hand, it is not treated as an instance of metalinguistic negation, it presents the same problems to a theory of presupposition projection as does (62b).

It is on the basis of considerations like these that I have argued that Horn's metalinguistic negation is not a natural category; it forces together phenomena that don't belong together. Hence, there is little hope for the position that presupposition denials should be accounted for in non-presuppositional terms. This is bad news for the satisfaction theory because

this theory cannot account for such cases, as we will see in the next chapter (§ 3.5). It is good news for the binding theory, which explains the preferred interpretations of the examples in (62) without further ado.

Before we leave this subject, at least for a while, let us consider why (62b) requires a marked intonation pattern, while (62a) does not. The answer to this question is not hard to find. In (62a) and (62b) the information that Harry doesn't have a wife is already in at the point at which the definite NP *Harry's wife* is evaluated, and therefore the issue of where the corresponding presupposition must be accommodated is settled right away. In (62b), by contrast, the hearer doesn't know for sure that the speaker intends this presupposition to be accommodated locally until the second sentence has been received. That is to say, the final interpretation of the first half of (62b) is partly dependent upon the interpretation of the second half. It is to signal this that the intonation contour required by denials like (62b) is a marked one, with a final rise within the negative clause, which the speaker uses to indicate that his contribution is not yet finished (cf. Horn 1989: 374; Ladd 1980: 145ff). In cases like (62b), by contrast, this signal is not required, and therefore this sentence can be pronounced with the more common falling intonation contour.

### 2.4.3 Bridging

Some presuppositions are harder to accommodate than others; for example, the presuppositions triggered by pronouns are rarely construed by way of accommodation. When produced 'out of the blue' the following will generally be felt to be infelicitous, presumably because it isn't clear what the speaker is referring to:

(65)     It's yellow.

By and large, a pronoun requires that an explicit antecedent be available, and as long as we ignore the possibility that antecedents are furnished by the context of discourse, this means that, generally speaking, a pronoun requires an antecedent *expression* for it to be interpretable. However, there are cases in which a pronoun can be used even in the absence of an antecedent expression. (65) might be the beginning of a riddle, for example:

(66)     It's yellow with black dots and two feet tall. What is that?

Here the only way to interpret the pronoun is by means of accommodation, obviously. The same holds, in my opinion, for what is sometimes called 'kataphora', that is, cases in which, informally speaking, a pronoun 'refers forward' to an antecedent that is yet to be formally introduced.

(67)     The doctor kept warning him but he never would listen, and now
         Harry's dead.

In this case, too, the personal pronoun *him* must be construed by way of
accommodation (*he* is an ordinary anaphor); all the hearer can do is set up a
reference marker representing a male individual. This representation is a
provisional one in the sense that it is not clear yet whom the speaker is
referring to. There are some non-trivial problems with this account, but these
are of no consequence to the main point that I want to make here, which is
that a pronoun that 'refers forward' must be construed by means of
accommodation. The phenomenon of forward reference will be further
discussed in § 4.3.

These examples were merely to show that sometimes even the
presuppositions triggered by pronouns can be accommodated. The cases
that I want to discuss in this section are in a sense intermediate ones. On the
one hand, it is not obvious that they are instances of accommodation. On the
other hand, they involve uses of pronouns that don't seem to require an
explicit linguistic antecedent. More accurately, these pronouns pick up
reference markers that are not made available by linguistic antecedents in
the standard way. The following examples are from Douloureux (1971):

(68)     a.     John bled so much *it* soaked through his bandage and stained
               his shirt.

         b.     When Little Johnny threw up, was there any pencil-eraser in
               *it*?

In both of these cases a pronoun occurs that is perfectly comprehensible
although it doesn't pick up an object that was explicitly introduced.[14]
Intuitively, the reason why such uses are possible is that, although the
intended referent has not been properly introduced, it can be inferred from
what has been said already, and is sufficiently salient. So these are instances
of binding, not accommodation, but they are different from the standard

---

[14] Some speakers of English will judge that such uses are sloppy and should be avoided, but that
doesn't seem to deter others from saying things like this. Note, furthermore, that the use of the
pronoun in (68b) is a functional one in the sense that it allows the speaker to avoid the word
*vomit* and its synonyms. I have heard such uses being called 'taboo pronouns', which I find a
happy expression, but I don't know who has coined this term. That the use of taboo pronouns is
not without risks is shown by the following example, whose source I cannot reveal:

When Fred was $\left\{\begin{array}{l} \text{decapitated} \\ \text{circumcised} \end{array}\right\}$, it was caught in a basket.

Examples like (68a, b) were much discussed in the seventies and early eighties, but seem to have
gone out of fashion since. For further observations and discussion, see Lakoff and Ross (1972),
Hirst (1981: 14–16), and Bosch (1983: 134–137).

cases of pronominal binding because the intended referents must be inferred.

In the case of pronouns, binding to inferred antecedents is a very constrained process, but it is none too clear what exactly the constraints are, as is shown by the following examples (from Lakoff and Ross 1972 and Hirst 1981):

(69)    a.    John became a guitarist because he thought that *it* was a beautiful instrument.

        b.    ?John became a flautist because he thought that *it* was a beautiful instrument.

        c.    ?John was a guitarist until he lost *it* on the subway.

Obviously, there are several factors that conspire to determine whether or not a pronoun can be used to pick up an inferred referent. The contrast between (69a) and (69b) suggests that phonological factors play a role: whereas the former contains the word *guitar* (in a way), the latter does not contain the word *flute*. But as the contrast between (69a) and (69c) shows, there are further factors involved, such as genericity, perhaps.

(68a, b) may be viewed as special cases of what Haviland and Clark (1974) have dubbed 'bridging'. To be sure, it is not common practice to subsume such examples under this notion; but then again this type of example is seldom discussed in the first place. Haviland and Clark introduce the notion of bridging for dealing with examples like the following:

(70)    Mary got some picnic supplies out of the car. The beer was warm.

The definite NP *the beer* does not have an overt antecedent, so the hearer must build an inferential bridge connecting the newly introduced referent with the preceding discourse. As far as I can see, there are no relevant differences between this example and the examples I discussed in the foregoing: in all these cases a presuppositional expression is used to refer to an object which was inferred from the previous discourse. This description also fits the following examples:

(71)    a.    Fred didn't forget to bring an umbrella. *It*'s in the hallway.

        b.    It's not true that Fred didn't bring an umbrella. *It*'s in the hallway.

In both cases it is clear what the intended referent of the pronoun is, although neither pronoun has a linguistic antecedent. The indefinite *an umbrella* really is a phantom antecedent, because it introduces its reference marker in a position that is not accessible from the position in which the pronoun occurs. So the pronoun cannot pick up a suitable reference marker unless one is inferred. In (71a) the first statement may be taken to *imply* that Fred brought

an umbrella, and this object is salient enough to be picked up by a pronoun; the same holds, mutatis mutandis, for (71b).

There is of course a difference between (71a) and (71b). The first statement in (71a) does not entail that Fred brought an umbrella; it merely implies this. The first statement in (71b), on the other hand, does entail that Fred brought an umbrella. However, I don't see that this difference makes much of a difference. Although the notion of bridging is often associated with non-monotonicity, defeasible inference, abduction, and so forth, there is no reason for supposing that bridging inferences are defeasible by definition. (71b) is just as much an instance of bridging as (71a) is.

In § 2.3 it was shown how the binding theory can deal with examples like (72a).

(72)    a.    Either Fred doesn't have a rabbit or Fred's rabbit is in hiding.
             (= (55))

        b.    Either Fred doesn't have a rabbit or *it* is in hiding.

The definite NP *Fred's rabbit* cannot be bound to the reference marker introduced by *a rabbit* in the first disjunct, which is inaccessible to it. On the account I suggested, this means that the pronoun must be construed by way of accommodation, and since global accommodation is excluded by Gricean constraints on interpretation, the presupposition must be accommodated locally. Unfortunately, this explanation does not carry over to (72b). If it were possible to accommodate the pronoun in (72a), we would get the following:

(73)    [x, z: Fred x, non-human z,
        [: ¬[y: rabbit y, x owns y]] ∨ [: in-hiding z]]

Here the presupposition triggered by the pronoun, i.e. [z: non-human z], is accommodated in the main DRS, and since this is compatible with the implicature that the speaker doesn't know if Fred has a rabbit, this reading should actually be preferred. But, of course, (73) is a most unlikely reading for (72b) to have.

Following Kamp and Reyle (1993: 187–190), I want to suggest that (72b) is a case of bridging (Kamp and Reyle don't use the term 'bridging', but that is what their proposal amounts to). On this view, (72b) is interpreted, in effect, as 'Either Fred doesn't have a rabbit or else he does have a rabbit and it is in hiding'. This inference is supported by the fact that, in general, disjunctive sentences receive an exclusive interpretation. A sentence of the form 'φ or ψ' is typically, though not necessarily, interpreted as 'φ or ψ, but not both'. It is this exclusive reading of *or* which allows the hearer to construe 'not-φ or ψ' as 'not-φ or (φ and ψ)'. It is perhaps still a moot question how the exclusive interpretation of *or* is to be accounted for, but whatever the right answer to

this question will turn out to be, it should be pointed out that the type of anaphora exemplified by (72b) is not a matter of logic alone. To see this, consider (74a), which is the disjunctive counterpart to (71a):

(74)     a.     Either Fred forgot to bring an umbrella, or it is in the hallway.

          b.     Fred didn't forget to bring an umbrella.

If the disjunction in (74a) is construed exclusively, the negation of the first disjunct may be assumed in the second, and once (74b) is given it is plausible to infer that Fred brought an umbrella, which is what enables the anaphoric pronoun to be interpreted. But (74b) does not entail that Fred brought an umbrella. So, (74a) is not relevantly different from (71a).

   If examples like (72b) and (74a) are to be treated along these lines, then there is no reason to withhold this treatment from (72a). Indeed, since there is every reason to expect that these sentences are interpreted the same way, I am forced to concede that (72a) is an instance of bridging, too. Nevertheless, I am reluctant to give up the accommodation account altogether, for two reasons. First, there are other examples of presupposition projection in disjunctions for which this is the only account that will work. The following type of case was first discussed by Liberman (1973):

(75)     Fred is either dating Barney's wife or his widow.

It will be clear that an explanation in terms of bridging is inappropriate here, and that an account in terms of accommodation is called for. A second, perhaps less compelling, consideration is that in cases like (72a) the accommodation and bridging accounts are not at odds with each other; they give the same results. It is possible, therefore, that a hearer may use either method to arrive at the intended reading. In the pronominal case, this is not so, because pronouns are generally difficult to interpret by way of accommodation, and besides accommodation will produce a different reading, as we have seen.

   It may be rare for a pronoun to be bound to an antecedent that was not explicitly introduced in the previous discourse, and must therefore be inferred. However, with other types of presupposition triggers bridging is much more common.

(76)     a.     If Wilma is married, then her husband is a lucky fellow.

          b.     If Fred doesn't want to come, Barney won't come either.

The definite NP *her husband* in (76a) does not have an antecedent expression, but of course if Wilma is married she will have a husband, so a suitable antecedent is easily inferred. Similarly, in (76b) the particle *either* triggers the presupposition that someone other than Barney will not come,

and although this is not explicitly given in the context in which the presupposition occurs it may be inferred from what is assumed in the antecedent of the conditional.

The following type of example was first discussed by Soames (1982):

(77)     If Harry has fallen into a depression again, his therapist will have a hard time getting him out of it.

There are (at least) two ways of interpreting this sentence. On the one hand, it may be read as implying that Harry has a therapist, who is going to find it difficult to cure Harry of his depression. On the other reading, the sentence may be paraphrased as follows:

(78)     If Harry has fallen into a depression again, then he will have to be treated by a therapist, who will have a hard time getting him out of his depression.

On the first reading, the presupposition triggered by *his therapist* is accommodated globally; on the second reading, it is bound to an antecedent derived by means of a bridging inference. As far as I can tell there isn't a strong preference for one reading or the other.

Is this then a counterexample against principle (B) of the binding theory, which says that binding is preferred to accommodation? Not really. It is just that since this principle was introduced we have extended the concept of binding. Recall that the formal notion of binding is to capture the insight that presupposed information is taken to be given in the discourse context. But, obviously, inferred antecedents are not immediately given in this sense. So if we allow for the possibility that presuppositions may be bound to antecedents that are inferred, we concede, in effect, that the distinction between binding and accommodation is not as clean as it seemed to be at first. Another way of making the same point is by noting that, instead of viewing bridging in terms of binding, we might just as well view it in terms of accommodation. That is, instead of allowing that a presupposition is bound to an inferred object, we might also have said that once a presupposition has been accommodated, it may be integrated into the discourse representation with the help of bridging inferences. The latter view meshes well with Heim's observation that a presupposition can almost never be accommodated without further ado (Heim 1982: 372–375, Heim 1992). A presupposition is presented as given information, and although a speaker may convey new information by presuppositional means, he cannot expect that his audience will accept this unless the presupposed information is unremarkable. It is common for people to own dogs (unfortunately), which is why I can introduce Barney's dog with a presuppositional device, as I just did. It is not common, in this country at least, for people to own yaks, so I shouldn't introduce Barney's yak the way I just did. Hence, if a presupposition is to be

accommodated, assumptions will have to be made about what is or is not remarkable, and this may be seen as (a form of) bridging.

So we have two views on bridging. A bridging inference either introduces an object for a presupposition to be bound to, or it helps to integrate an accommodated presupposition into the discourse representation. Both views have something to recommend them. Sometimes, an inferred antecedent can be as salient as if it had been introduced explicitly. This applies to (68a, b), for example. But there are also cases in which the inference is less obvious. The first time I read Haviland and Clark's (1974) famous example of bridging, reproduced above as (70), I had a slight difficulty understanding the second sentence, because at the time my stereotype of picnics did not make provision for alcoholic beverages of any kind. Despite this cultural barrier, I was able to interpret *the beer* as intended by Haviland and Clark. I did this, apparently, by first accommodating the presupposition triggered by the definite NP and then trying to figure out how it might relate to the preceding discourse.

From the foregoing discussion one might draw the depressing conclusion that the notions of binding and accommodation we have developed are inadequate, because they ignore the crucial role that world knowledge and plausibility inferences play in presupposition projection. But I prefer a more uplifting view. The binding theory tries to isolate the semantic and pragmatic principles governing presupposition projection from factors having to do with world knowledge and plausibility. Of course, this project presupposes that there *are* principles governing presupposition projection, and that these *can* be separated from factors having to do with world knowledge and plausibility. But these premises require hardly any justification. What is called the projection problem for presupposition is an intricate pattern of empirical regularities with systematic exceptions, which can be described without reference to world knowledge and plausibility inferences, as I have demonstrated in § 1.1. It is of course conceivable that this pattern can only be explained by a theory which holds that presupposition projection is inextricably tied up with speakers' beliefs and guesses. But this possibility seems quite remote. We can grant the relevance of world knowledge and plausibility inferences, and that they may interact with principles governing presupposition projection; but this doesn't obviate these principles, or the notions of binding and accommodation in terms of which they are formulated.

None of this is to imply that it would be difficult to incorporate bridging into the binding theory. It is just that I don't see that such an exercise would be particularly useful. Being a version of DRT, the binding theory imposes various constraints on the set of reference markers a given marker $u$ may be identified with. But it doesn't specify *which* of these $u$ must be bound to: the projection mechanism takes as given a (possibly empty) set of bindings for

any given presupposition. In other words, although one of the subjects of the binding theory is anaphora, it is not a theory of anaphora *resolution*. This division of labour is uncontroversial: it is assumed in virtually all linguistic and philosophical work on anaphoric reference. But of course the same division of labour should apply when we want deal with bridging phenomena.

It would be easy enough to generalize the binding theory by taking as given, for any presupposition, not only a set of bindings but also a set of bridges; such bridges would take the form of DRSs, of course. In this type of theory, bridges would have the same status as bindings: the theory would impose some restrictions upon bridging, and it would make predictions about where a given bridge must be inserted in the discourse representation. It is plausible to assume, for example, that binding would be preferred to bridging, which in its turn would be preferred to accommodation. But such a theory would have little of interest to say about bridging, for the same reason that most linguistic theories of anaphora have little of interest to say about anaphora resolution. Something like this can be done, surely, but for reasons explained in the foregoing I will not do it.

### 2.4.4 Intermediate accommodation and domain restriction

If a presupposition must be accommodated and cannot be accommodated in the main DRS, the binding theory predicts that there is a preference for accommodating it further down in the chain of DRSs linking the presupposition's home DRS with the main DRS. Thus it may happen that a presupposition is preferably accommodated in a DRS that is neither its home DRS nor the main DRS. Let us refer to such cases as instances of intermediate (as opposed to local or global) accommodation.

(79)     Maybe she has found out that her husband is deceiving her.

In (79), the definite NP *her husband* induces the presupposition that 'she' has a husband; this presupposition is triggered within a belief context that, in its turn, is embedded under a modal operator. In the absence of a suitable antecedent this presupposition will by default be accommodated in the main DRS, which yields a reading that may be paraphrased as follows: 'She has a husband$_i$ and maybe she has found out that he$_i$ is deceiving her.' If for some reason this reading is dispreferred, then the second option is intermediate accommodation, which results in the following: 'Maybe (she has a husband$_i$ and has found out that he$_i$ is deceiving her)'. This would be the preferred reading in a context like the following.[15] A woman is sitting on a bench in the

---

[15] This type of scenario was invented by Fauconnier (1985). See Kay (1992) for further discussion.

park, showing sundry signs of grief (moaning, tears, wringing of hands). Fred
and Barney pass her by, and it is mutually clear to them that neither knows
the woman. Then Fred might say (79) to Barney, in an attempt to explain the
woman's obviously deplorable state. In this scenario, the intended reading of
(79) would require intermediate accommodation.

   Another variety of intermediate accommodation occurs when a
presupposition $\chi$ is triggered in the scope of a quantifying expression $\alpha$ and
$\chi$ contains a reference marker bound by $\alpha$. The binding theory predicts that
in such an event global accommodation is excluded (because the resulting
DRS would not be a proper one), and on the next-preferred reading $\chi$
restricts $\alpha$'s domain. The following is a case in point:

(80)     Everyone of my friends has sold his copy of 'Lolita'.

This will ordinarily interpreted as saying that everyone of my friends who
had a copy of 'Lolita' sold it. On this reading, an occasional friend who didn't
have a copy of 'Lolita', and therefore couldn't sell it either, does not falsify
(80).[16] To show in some detail how the binding theory accounts for this
reading, I will employ Kamp and Reyle's (1993) 'duplex conditions'. This is
not because I want to advocate this proposal, but because this is as good an
opportunity as any for explaining why I will not adopt it later on.

(81)     a.   [: [x: friend x]⟨every x⟩[u̲, v̲: <u>Lolita v, u owned v</u>, x sold v]]

         b.   [: [x, u: u = x, friend x]⟨every x⟩
                 [u, v̲: <u>Lolita v, u owned v</u>, x sold v]]

         c.   [: [x: friend x]⟨every x⟩[v̲: <u>Lolita v, x owned v</u>, x sold v]]

         d.   [: [x, v: friend x, Lolita v, x owned v]⟨every x⟩[: x sold v]]

(81a) is the semantic representation of (80) in which only the two
presuppositions triggered by *his copy of 'Lolita'* remain to be processed.
(81a) contains a single duplex condition of the form '$\varphi$⟨Q  *u*⟩$\psi$', where $\varphi$ and
$\psi$ are DRSs, Q is a quantifier, and *u* is a reference marker. I take it that the
intended interpretation of this structure is transparent enough, and will not
discuss it in detail.[17] The definite NP *his copy of 'Lolita'* triggers the two-part
presupposition that (i) there is an individual u such that (ii) v is the copy of
'Lolita' that u owned. The first presupposition is bound to the reference

---

[16] In § 1.4, I already discussed some examples of this type, which I repeat here for convenience:

   Many professors sold their Coca Cola shares.
   Theo usually stops drinking before he gets sick.
   Everyone should leave his camera at the information desk.

[17] See Kamp and Reyle (1993: Chapter 4) and Geurts and van der Sandt (1999) for further
discussion.

marker x in the domain of the quantifier, as shown in (81b), which is equivalent to (81c). The second presupposition cannot be bound and must therefore be accommodated. Accommodation in the principal DRS is not possible because this presupposition contains a reference marker, i.e. x, which is introduced in the domain of the quantifier, and so the binding theory predicts that accommodation in the restrictor is the next-preferred option, and we obtain the DRS in (81d). The reading thus obtained is roughly (but *only* roughly, as we will presently see) paraphrased by (82):

(82)    Everyone of my friends who owned a copy of 'Lolita' sold it.

I think this is a gratifying result, but it has been argued by Beaver (1994) that this prediction is false. Beaver's objection is based upon the supposition that, according to the binding theory, there is no relevant difference at all between (80) and (82). If this were so, then the following contrast would become a problem:

(83)    a.    ?Hardly any of my friends ever owned a copy of 'Lolita', and everyone of my friends has sold his copy of 'Lolita'.

        b.    Hardly any of my friends ever owned a copy of 'Lolita', and everyone of my friends who owned a copy of 'Lolita' sold it.

Whereas (83b) is a fairly normal thing to say, (83a) sounds decidedly odd. However, the binding theory predicts, according to Beaver, that (80) and (82) are equivalent, so it cannot account for the contrast observed in (83). More generally, Beaver charges, the prediction that a presuppositional expression in the scope of a quantifier can give rise to domain restriction is incorrect.

It is not hard to see what is wrong with Beaver's objection. There is an obvious difference between (80) and (82), which happens to be of a presuppositional nature. In (80) the quantified subject NP triggers the presupposition that a set of friends is given. The corresponding presupposition triggered in (82) is more specific: it presents as given a set of friends who own a copy of 'Lolita'. If the binding theory had predicted that the same presuppositions are triggered in (80) and (82), Beaver would have had a point. But since it doesn't, he hasn't.

This suffices to refute Beaver's objection, but of course there is more to say about the difference between (80) and (82), and how this difference causes the contrast in acceptability between (83a) and (83b). We must somehow account for the fact that (80) and (82) have different presuppositions, and the problem with Kamp and Reyle's version of DRT is that it doesn't provide the means for doing this in a straightforward way. The presupposition triggered by the quantifier in (80) may be paraphrased as 'the speaker's friends'. It may be thought of as the set of individuals which satisfy the

restrictor of the duplex condition in (81a). Using Kamp and Reyle's sum operator, this could be represented as follows:

(84)      [X: X = Σx[x: friend x],
                [x: friend x]⟨every x⟩[u, v: Lolita v, u owned v, x sold v]]

Here 'Σuφ' stands for the set of *u*'s which satisfy φ. This procedure is a bit clumsy because we have to copy material from the quantifier's restrictor into the description associated with the set represented by X; and it isn't fully adequate either. For if the presupposition triggered by *his copy of 'Lolita'* is accommodated in the restrictor, as the binding theory predicts, it should be copied over to the sum description, too. That is, the reading we eventually want to assign to (80) is the following:

(85)      [X: X = Σx[x, v: friend x, Lolita v, x owned v],
                [x, v: friend x, Lolita v, x owned v]⟨every x⟩[: x sold v]]

This should be the interpretation of (82), too. The difference between (80) and (82) consists in the way this interpretation comes about. In the latter case, the set description 'Σx[x, v: friend x, Lolita v, x owned v]' is generated *in toto*; it represents the presupposition triggered by the subject term. In the case of (80) this description is built in two steps. We first have 'Σx[x: friend x]', which represents the presupposition triggered by *everyone of my friends*, which is then restricted by the presupposition triggered by the possessive NP. This procedural difference can be expressed as follows. A speaker who utters (82) presupposes the set of individuals that are his friends and own copies of 'Lolita'. By contrast, a speaker who utters (80) presupposes the set of individuals that are his friends, *and* he presupposes that they own copies of 'Lolita'. These are different things altogether, as the contrast between (83a) and (83b) shows.

Having dealt with Beaver's observations, let me conclude this section with a remark on Kamp and Reyle's treatment of quantification. The representational machinery introduced by Kamp and Reyle is quite powerful, and it will be hard to prove that there are things it cannot do in principle. However, Kamp and Reyle's theory of quantification can hardly be claimed to be elegant. Even for the fragment that Kamp and Reyle are concerned with, it is often necessary to copy material from one corner in a DRS to another, which is something one should wish to avoid. This only becomes worse, as we have seen, if we want to take into account the presuppositions triggered by quantifiers.

Informally speaking, one would wish that the DRS representation of a quantifier domain could be moved about. If a quantified sentence is represented by a condition of the form 'φ⟨Q *u*⟩ψ', and φ is presupposed, then φ should be the kind of object that can be bound or accommodated. Since, in Kamp and Reyle's version of DRT, φ is a DRS, this is not possible.

In the version of DRT introduced in Chapter 5, however, quantifier domains have the same status as reference markers, which makes for a much simpler treatment of the presuppositions triggered by quantified expressions.

## 2.5  Further remarks on binding and accommodation

The central concepts of the binding theory are binding and accommodation. Technically speaking, these concepts are simple enough. If a presupposition $\chi$ is bound in a DRS $\varphi$, the content of $\chi$ and a set of equations is merged into $\varphi$. Accommodation is a special case of this: if $\chi$ is accommodated the set of equations is empty. The binding theory consists of a set of principles that constrain the operation of these mechanisms.

We have already seen that the binding theory does not capture, nor is designed to capture, presupposition projection in all its facets. DRT is, inter alia, a theory of anaphora which abstracts away from the role played by world knowledge and plausibility inferences in the resolution of anaphoric pronouns. The binding theory is an extension of DRT whose aim is to account for the pragmatic principles governing presupposition projection, while remaining aloof from world knowledge and plausibility. These factors are acknowledged (or kept out, depending how you look at it) by the proviso that the principles of the binding theory only hold ceteris paribus. If all things are not equal, these principles may be defeated.

I have been careful to avoid the claim that a presupposition is either construed anaphorically or else accommodated. In this respect I disagree with van der Sandt, although the disagreement merely concerns a terminological matter. Van der Sandt (1992: 341) claims that all presuppositions are anaphors, and in doing so inflates the traditional concept of anaphora beyond recognition. In my vocabulary, a theory of presupposition *subsumes* the traditional concept of anaphora, and the following, it seems to me, is a rather good approximation of that concept:

(86)    *Anaphora*
        A presuppositional expression functions anaphorically whenever it (more accurately: its semantic correlate) is bound to some discourse entity which is at the focus of attention.

This is not intended as a definition, of course, for it contains at least one term whose meaning is somewhat vague, but my main concern is to show how anaphora fits into the general picture. Note, to begin with, that this quasi-definition doesn't say what types of expressions are anaphors; rather, it states what it means for a given token to function anaphorically. I take it that some pronouns will virtually always function anaphorically in this

sense, and those which do might be classified as anaphors in a derivative sense. But the same will hold true of semantically attenuate full NPs, like *the thing* or *the animal*.

Given the right kind of context, even the poorest pronouns can be interpreted by way of accommodation (§ 2.4.3). But such cases are exceptional: pronouns and other semantically attenuate definites are generally anaphoric. Many definite NPs are not like this. Descriptions like *the German president* and *the moon*, to name only two examples, are specific enough not to be restricted for their reference to those entities which are in focus. It might be thought that at least some of these expressions are problematic for the binding theory. For it seems that the theory predicts that in a discourse in which the moon (or a moon) has not been mentioned yet, the definite NP *the moon* would have to be construed via accommodation. But this conclusion is unacceptable, for surely *the moon* will normally be used to refer to the moon. There is no deep problem here, however, for as we have already seen in § 2.4.3, the universe of a DRS need not be restricted to objects which have been introduced explicitly and by linguistic means. Since a DRS is a representation of the speaker's commitment slate, it is natural to assume that it may also contain information which derives from other sources. Most speakers may be assumed to know that there is a moon, so we can maintain after all that, as a rule, the definite NP *the moon* is construed by way of binding rather than accommodation.

Some presuppositions are easier to accommodate than others. While pronominal presuppositions are usually bound, the presuppositions triggered by certain factive verbs or definite NPs are quite readily accommodated. Van der Sandt (1992) suggests that such differences correlate with the descriptive richness of presuppositions: pronouns are semantically attenuate expressions, and therefore must be bound, whereas the presuppositions triggered by factives may be quite specific, and therefore can be accommodated. This explanation is a plausible one, but it is incomplete. For instance, it seems that the presuppositions triggered by *too* may be as specific as factive presuppositions, but nonetheless the latter are easily accommodated, while the former are not. Compare:

(87)     a.     Fred doesn't know that someone tried to steal his lawn mower.

         b.     Barney tried to steal Fred's lawn mower too.

The factive in (87a) triggers the presupposition that someone tried to steal Fred's lawn mower. (87b) has almost the same presupposition; assuming that the focus is on *Barney*, the presupposition triggered is that someone other than Barney tried to steal Fred's lawn mower. This is, if anything, more specific, but nonetheless the first presupposition is easily accommodated, while the second one must be bound. Or compare:
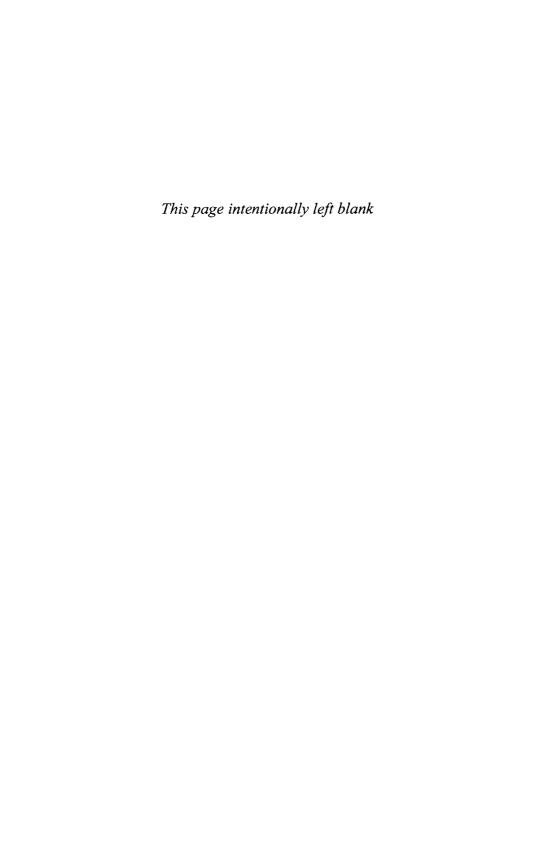
(88)  a.  Did you know

b.  Isn't it a shame that $\Bigg\}$ Wilma is pregnant again?

(88a) and (88b) presuppose the same thing, viz. that Wilma is pregnant again. But obviously this presupposition is much more readily accommodated in the former case than in the latter. Apparently, it is not just the content of a presupposition that determines how easily the presupposition is accommodated, and therefore van der Sandt's conjecture doesn't provide a full explanation of why some presuppositions are harder to accommodate than others. Unfortunately, I am still at a loss to see how a better explanation might go.

It bears emphasizing that there are probably no triggering expressions whose presuppositions must always be bound. We have seen that even pronouns can be construed by way of accommodation, and if *it* can be accommodated, chances that some other presuppositional expressions must always be bound would appear to be remote. I know of just a single possible exception to this claim. Inspired by Kripke's observations on *too*, as reported by Soames (1989), a number of authors have proposed that *too* is inherently anaphoric. According to Heim (1992), for example, *too* introduces an anaphoric element which must be coindexed with a earlier expression. But it is not true, as is often suggested, that *too* always requires an explicit antecedent. The day after India detonated its first series of nuclear bombs, I read a newspaper column that began as follows:

(89)    We now know that India has nuclear arms, too.

Even in the absence of an antecedent expression it is reasonably clear what the *too* refers to. At any rate, this sentence is fully acceptable, and although it is a matter of debate whether this is an instance of accommodation or not, it demonstrates that the presupposition triggered by *too* need not always be given in the preceding discourse.

*This page intentionally left blank*

CHAPTER 3

# The satisfaction theory

Currently, the most popular accounts of presupposition projection are the binding theory and the satisfaction theory. Prima facie, these two theories are rather similar, and it has even been suggested that they are essentially equivalent (Zeevat 1991, Heim 1992). In this chapter it will be shown, however, that there are real and important differences between these two accounts.

The ideas underlying the satisfaction theory are often traced back to Stalnaker (1973, 1974) but properly seen Karttunen (1974) is the theory's founding father; Stalnaker's views are more in line with the binding account. Karttunen's ideas were shaped into an explicit theory by Heim (1983), which was subsequently taken up by Chierchia and McConnell-Ginet (1990), van Eijck (1993), Beaver (1992, 1993, 1994), Heim (1992), Krahmer (1993), Krifka (1993), and Chierchia (1995). In the following I will mainly criticize the satisfaction theory on the grounds that its predictions are too weak. I feel justified in referring to *the* satisfaction theory of presupposition because my objections apply to all varieties that I know of. In fact, they apply to many other theories besides, such as Karttunen and Peters's (1979) and some logical presupposition theories (for example, van Fraassen 1969), but these fall outside my target area.[1]

The satisfaction theory has been attacked before, especially by Gazdar (1979) and van der Sandt (1988), but up to a point I concur with Soames's (1982) and Heim's (1983) judgment that these criticisms have been inconclusive. Furthermore, many of the arguments that have been launched against the satisfaction theory have become outdated because they are based upon assumptions that aren't generally accepted anymore, such as the idea that presupposition projection should be explained by a single set of rules or principles: as it is now widely agreed that any viable explanation of the presuppositional facts will be a hybrid one, at least some of the arguments that were aired in the seventies have lost their potency. The

---

[1] Conceptually, the theory proposed by Karttunen and Peters (1979) is quite different from Karttunen's 1974 theory (and much more in the spirit of his 1973a theory), which I consider to be the first representative of the satisfaction theory. Empirically, however, the two theories are equivalent, and consequently they are often treated on a par, for example, by Gazdar (1979).

objections presented in this chapter are mostly new ones, but it should be noted that they are aimed at the same weaknesses that Gazdar and others have complained about, and that the main problems have been known for more than two decades.

This chapter is organized as follows. In the first section, § 3.1, the satisfaction theory is introduced, and in § 3.2 it is shown that the theory's predictions are too weak. In particular, the satisfaction theory often predicts presuppositions of the form $\varphi \to \psi$ where the intuitively perceived presupposition is simply $\psi$. I call this the 'proviso problem'. The proviso problem has been tackled in the literature by assuming that there is some process which, in some way or other, allows the hearer to conclude from a conditional presupposition $\varphi \to \psi$ that the speaker in fact wants to convey that $\psi$ is true. Thus the inference that the speaker is committed to the truth of $\psi$ is obtained in two steps, and I argue that this by itself is already a questionable set-up. In §§ 3.3 and 3.4, I examine the two proposals for implementing this two-step model which have been suggested in the literature, and find them both wanting. Moreover, I present evidence that strongly suggests that no such model could ever work. Some further problems for the satisfaction theory are discussed in § 3.5. Finally, in § 3.6, I compare the satisfaction theory with the binding theory, and try to explain why the former but not the latter runs into the problems discussed in the preceding sections.

## 3.1 The theory

Like the binding theory, the satisfaction theory draws its main inspiration from Stalnaker's insight that, in general, an utterance changes the context in several consecutive steps, between which local (or intermediate) contexts are constructed, and that this process of context change is the key to the solution of the projection problem. But the satisfaction theory departs from Stalnaker's original ideas in a quite radical manner, in that it adopts the position that *sentences* (not utterances) change contexts, and that the compositional semantics of an expression (not only of sentences) may be defined in terms of what it does to its context. In a word, the satisfaction theory requires a dynamic semantics, which defines meaning in terms of context change, and the theory claims that once we have such a semantics, the projection problem will be solved, too. This is a clean break with Stalnaker's views. When Stalnaker speaks of context change, he has pragmatic processes in mind, and his notion of presupposition is a thoroughly pragmatic one, too. The satisfaction theory is an attempt to recast Stalnaker's pragmatic theory in a semantical mould. To some readers, this may seem like an attempt to square the circle, and it will gradually become clear that this analogy is not entirely out of place.

According to satisfaction theory, a context may be viewed as an information pool, and sentences as operations on contexts, thus understood. For example, the context that results from uttering a simple sentence $\varphi$ will contain all the information that the initial context contained, plus the information that $\varphi$ is true. Now take a simple sentence that contains a presupposition-inducing expression, which (as in the preceding chapters) we write as $\varphi\{\chi\}$, where $\chi$ is the presupposition that is triggered in $\varphi$. According to the satisfaction theory, the presupposition $\chi$ requires that the context which is being incremented with its carrier sentence $\varphi$ already contain the information that $\chi$ is true. More succinctly: if $\varphi\{\chi\}$ is to be added to c, then c must satisfy $\chi$. If this requirement isn't met, then c is not an appropriate context for $\varphi\{\chi\}$. Heim would say that $\varphi\{\chi\}$ is 'not defined' in c or that c doesn't 'admit' $\varphi\{\chi\}$,[2] and Karttunen (1974) that c doesn't 'satisfy-the-presuppositions-of' $\varphi\{\chi\}$. In any case, the intuition is that c cannot without further ado be incremented with $\varphi\{\chi\}$ unless it follows from c that $\chi$ is true. This is not to imply, however, that $\varphi\{\chi\}$ just cannot be appropriately uttered if c doesn't satisfy $\chi$: officially c must satisfy $\chi$, but if it doesn't, a cooperative hearer will generally be prepared to revise c so as to make it satisfy $\chi$, before he goes on to process $\varphi$. That is, $\chi$ will be accommodated in c, and $\varphi$ interpreted in the resulting context. In this way, a presupposition may carry information that is strictly speaking new, although under the pretence that it is already given in the context. This is Stalnaker's account, but with the non-trivial difference that sentences have taken the place of utterances.

Thus far we have confined our attention to simple sentences and their presuppositions, but the satisfaction theory's crucial tenet is that the same story can be told about embedded sentences, because they, too, may be viewed as context-change devices. To illustrate this idea, let us consider how the sentences of a simple propositional language can be modeled as context-change devices. We adopt Veltman's (1996) notation and let $c[\varphi]$ stand for the result of incrementing c with $\varphi$. A context is simply a subset of a given set of possible worlds W, and a sentence $\varphi$ will reduce the current context c to those worlds in c that make $\varphi$ true. If $\varphi$ is a simple sentence, this is a primitive operation, and the context change brought about by a complex sentence is defined in terms of the contextual effects of its parts, as follows:

(1)    a.    $c[\varphi] = \{w \in c \mid \varphi \text{ is true in } w\}$, if $\varphi$ is a simple sentence

        b.    $c[\neg\varphi] = c - c[\varphi]$

        c.    $c[\varphi \vee \psi] = c[\varphi] \cup c[\psi]$

        d.    $c[\varphi \wedge \psi] = c[\varphi][\psi]$

---

[2]  The term 'admit' was suggested, though not used, for this purpose by Karttunen (1974).

   e.    $c[\varphi \rightarrow \psi] = c - (c[\varphi] - c[\varphi][\psi])$

For any sentence $\varphi$, $[\varphi]$ is a total function whose domain is pow(W). If $\varphi$ is a simple sentence, then this function reduces its input context c to that subset of c in which $\varphi$ is true. Negation is in effect set complementation: in order to compute $c[\neg\varphi]$ we first compute $c[\varphi]$, and subtract the result from c. Disjunction is defined as set union, as one might expect, but a conjunction is processed from left to right: first c is updated with $\varphi$, and then the resulting context is updated with $\psi$. (From a logical point of view, this ordering has no effects, yet, but it will have when we extend the language in a moment.) The definition of the conditional may be somewhat opaque at first sight, but what it gives us is the material implication:

$$c[\varphi \rightarrow \psi] = c - (c[\varphi] - c[\varphi][\psi]) = c - c[\varphi][\neg\psi] = c - c[\varphi \wedge \neg\psi] = c[\neg(\varphi \wedge \neg\psi)]$$

The semantics in (1) extends the notion of context change to embedded sentences: in order to interpret a complex sentence $\varphi$ in c, each sentence embedded in $\varphi$ must be interpreted in its own local context, which may but need not coincide with c. To compute $c[\neg\varphi]$, for example, we have to evaluate $c[\varphi]$ as an auxiliary computation, and thus the local context of $\varphi$ as it occurs in $\neg\varphi$ is the same as that of its matrix. From (1c) it follows that the members of a disjunction, too, must be evaluated in the same context in which the matrix occurs. Conjunctions are processed in an incremental fashion, and therefore, if c is updated with $\varphi \wedge \psi$, then c is the local context of $\varphi$, while the local context of $\psi$ is $c[\varphi]$; the same holds for conditionals. In general: if c is updated with a sentence $\xi$, and $\xi$ contains a sentence $\xi'$, then the local context of $\xi'$ is given by an expression of the form $c[\zeta_1] \dots [\zeta_n]$, with $n \geq 0$; this will always be a subset of c. The notion of local context, thus characterized, is at the heart of the satisfaction theory's understanding of presupposition, and will play a central role in the following.

   We now extend our object language with a syntactic device for representing presuppositions that we have been using already: if $\varphi$ and $\psi$ are simple sentences, then $\varphi\{\psi\}$ is a sentence, whose intuitive meaning is that $\psi$ is a presupposition triggered in $\varphi$. Restricting this sentence-forming operation to simple sentences is a natural thing to do as far as the carrier sentence is concerned, but this restriction also excludes complex presuppositions, which is less realistic. However, since this choice simplifies matters without distorting the main issues, we will adopt it anyway. Let us say that a context c satisfies $\varphi$ iff $c[\varphi] = c$. Then we can extend our semantics as follows:

(1)    f.    $c[\varphi\{\psi\}] = c[\varphi]$, if c satisfies $\psi$; undefined otherwise

Thus our context-change semantics for propositional logic becomes a partial one, with the following definedness conditions for the connectives:

(2)    a.    $c[\neg\varphi]$ is defined iff $c[\varphi]$ is defined

b.   $c[\varphi \vee \psi]$ is defined iff $c[\varphi]$ is defined and $c[\psi]$ is defined

c.   $c[\varphi \wedge \psi]$ is defined iff $c[\varphi]$ is defined and $c[\varphi][\psi]$ is defined

d.   $c[\varphi \rightarrow \psi]$ is defined iff $c[\varphi]$ is defined and $c[\varphi][\psi]$ is defined

Once we have adopted (1f), conjunction ceases to be commutative; its definedness conditions mirror those of the conditional. The order of the members of a disjunction, on the other hand, remains arbitrary.

The definedness conditions in (2) follow from the semantic definitions in (1), and the predictions that the satisfaction theory makes about presupposition projection follow from these definedness conditions. Presuppositions are contextual requirements: if $\varphi$ presupposes $\chi$, then $c[\varphi]$ isn't defined unless c satisfies $\chi$. We can therefore define the notion of presupposition as follows:

(3)      $\varphi$ presupposes $\chi$ iff for all c, $c[\varphi]$ is not defined unless c satisfies $\chi$

Thus negation is predicted to be a hole, to use Karttunen's (1973a) term: $c[\neg\varphi]$ is defined iff $c[\varphi]$ is defined, and therefore the presuppositions of $\neg\varphi$ are the same as those of $\varphi$. I take it that this prediction is essentially correct, and the same holds for what the theory says about the projection characteristics of disjunctions. As (2b) indicates, it is predicted that a disjunction will inherit the presuppositions of its parts. In the majority of cases this is surely correct, but prima facie this prediction causes problems with the following types of example, which were discussed at some length in the previous chapter:

(4)      a.   Either Fred keeps his Chevrolet hidden somewhere or he doesn't have one in the first place.

         b.   Either Fred doesn't have a Chevrolet or he keeps his Chevrolet hidden somewhere.

(5)      Barney has either stopped smoking or he has begun to smoke (I forgot which).

As it stands, the theory predicts that (4a) and (4b) presuppose that Fred owns a Chevrolet, and that (5) has an inconsistent set of presuppositions, namely that Barney used to smoke and didn't smoke before the utterance time. But although these are serious problems, (4) and (5) are not straightforward counterexamples, because the satisfaction theorist can use local accommodation to explain these cases. Let us see how such an expanation would go in the present framework.

We said that, as a rule, a hearer will be prepared to accommodate presuppositions that are not satisfied by the current context. However, within a dynamic framework this informal description requires some sharpening:

we have to specify what is meant by 'the context' in which a presupposition arises. To illustrate, suppose that a sentence of the form $\varphi\{\chi\} \vee \psi$ is added to c (note that (4a) is of this form), and that c doesn't satisfy $\chi$. In this situation, $\chi$ might be accommodated in c — i.e. $\chi$ might be added to c as if it had been asserted in the foregoing. In this case $\chi$ is accommodated *globally*. However, it is also possible to only temporarily accommodate $\chi$ in c to allow for the intermediate evaluation of $c[\varphi\{\chi\}]$, as required by (1c). In this case $\chi$ is accommodated *locally*: $\chi$ is added to a copy of c, so to speak, and the global context itself remains untouched. In order to compute $c[\varphi\{\chi\} \vee \psi]$, we must evaluate $c[\varphi\{\chi\}]$ as an auxiliary computation, and local accommodation simply means that we take $c[\chi][\varphi\{\chi\}]$ instead of $c[\varphi\{\chi\}]$. Thus, if we opt for local accommodation, $c[\varphi\{\chi\} \vee \psi] = c[\chi][\varphi\{\chi\}] \cup c[\psi]$. For (4a), this means that we obtain the following reading: 'Either (Fred owns a Chevrolet and he keeps it hidden somewhere) or he doesn't own a Chevrolet,' which is what we want this sentence to mean. Similar remarks apply to (4b) and (5). So these examples can be accounted for if we assume that the presuppositions that initially appeared to be problematic are accommodated locally.[3]

It is obvious that, without further provisions, the distinction between global and local accommodation would give rise to the prediction that presuppositional expressions cause systematic ambiguities, a prediction that is not borne out by the facts. For example, if the following sentence (which is of the form $\neg\varphi\{\chi\}$) is uttered in a context which doesn't satisfy the proposition that it is raining,

(6)    Barney doesn't know that it is raining.

then the audience might either globally or locally accommodate the presupposition that it is raining. In the former case, the sentence would be construed as: 'It is raining and Barney doesn't know that it is raining'; in the latter, as: 'It isn't the case that it is raining and that Barney knows that it is raining.' Arguably, (6) allows for both construals, but it is evident that the first is strongly preferred. It must be assumed, therefore, that global accommodation is the default option.

In § 2.4.3, I argued that the accommodation analysis of examples like (4b) is unsatisfactory, because it doesn't extend to pronominal cases like (7), which obviously should be treated the same way as (4b).

---

[3] Incidentally, it is not at all easy to see how accommodation could be implemented within this framework. To be sure, it is intuitively clear what accommodation does as long as one is talking *about* a semantics along the lines of (1), but it is by no means trivial to integrate the notion of accommodation into such a semantics. See however Beaver (1992, 1993) for an account of global accommodation. To my knowledge, no formal theory of local accommodation has been proposed so far. Problems of implementation aside, it will be argued in § 3.5 that the satisfaction theory's account of local accommodation is untenable.

(7)     Either Fred doesn't have a Chevrolet or he keeps *it* hidden
        somewhere.

This is actually a big problem for the satisfaction theory, because it cannot
offer a unified treatment of presupposition and anaphora. We have seen that
there are systematic parallels between anaphoric and presuppositional
phenomena, which are accounted for by the binding theory. If the
satisfaction theory were true, it would seem that these parallels are just a
matter of chance. This is unsettling, to the say the least, but I will not pursue
this line of argument, because I do not merely want to argue that the binding
theory is better than the satisfaction theory. What I want to show, rather, is
that the satisfaction is wrong because it fails to account even for the simplest
presuppositional facts. In the next chapter I will return to examples like (7),
and consider how they might be treated in a dynamic semantics framework.

## 3.2  The proviso problem

The principal problem with the satisfaction theory is not that it is too strong,
as (4) and (5) initially seemed to suggest, but rather that it is too weak. This
defect of the theory, which I shall argue is beyond repair, manifests itself with
conjunctions and conditionals, and it is to these that we now turn. Given the
definedness conditions in (2c) and (2d), the theory predicts that conjunctions
and conditionals have the same presuppositional behaviour. Let us look at
conjunctions first. A sentence of the form $\varphi\{\chi\} \wedge \psi$ is predicted to presuppose
that $\chi$, which is correct: (8a) presupposes that (8b) is true.

(8)     a.     Fred's wife hates sonnets and so does his manager.

        b.     Fred has a wife.

However, if we take a sentence of the form $\varphi \wedge \psi\{\chi\}$, the presuppositional
requirement is *not* that the context satisfy $\chi$. For, $c[\varphi \wedge \psi\{\chi\}]$ is defined iff
$c[\varphi]$ is defined and $c[\varphi]$ satisfies $\chi$, and it is easily proved that this is the same
as requiring that $c[\varphi]$ be defined and c satisfy $\varphi \to \chi$. In other words, it is
necessary as well as sufficient for $\chi$ to be satisfied in its local context $c[\varphi]$ that
c satisfy $\varphi \to \chi$, and thus $\varphi \to \psi\{\chi\}$ is predicted to presuppose that $\varphi \to \chi$,
instead of $\chi$, as one would expect. For example, (9a) is predicted to
presuppose merely that (9b) is true.

(9)     a.     Fred hates sonnets and so does his wife.

        b.     Fred hates sonnets $\to$ Fred has a wife

Nonetheless, we would infer from (9a) that Fred has a wife, just as we would
from (8a), but this inference is easily accounted for, because if (9a) is

updated to a context that meets the requirement that (9b) is true, then the resulting context will satisfy the proposition that Fred has a wife, via modus ponens. Thus both sentences will license the inference that (8b) is true, but only (8a) presupposes (8b); (9a) merely presupposes (9b). Besides, one would probably want to say that (8a) and (9a) entail (8b) anyhow, so although it seems a bit peculiar that these sentences should have different presuppositions, the predictions we get are otherwise unobjectionable.

Things take on a different aspect, though, if we embed (8a) and (9a) under a negation operator or in a modal context. First, let us have a brief look at the type of example for which the satisfaction theory gives the correct predictions:

(10)    It is $\left\{\begin{array}{l}\text{posssible}\\\text{not true}\end{array}\right\}$ that Fred has a wife and that his wife hates sonnets.

Negation, as we have seen, is transparent to presuppositions: $\neg\varphi$ is defined iff $\varphi$ is, and I assume here that the same would hold for modal operators such as *possible*, although officially our fragment doesn't contain any. Thus (10) is predicted to impose upon its context the same presuppositional requirements as the embedded conjunction does, and if we may pretend that this sentence contains just a single presupposition-inducing expression, i.e. the possessive NP, the set of its presuppositional requirements is empty. For the embedded conjunction is of the form $\varphi \wedge \psi\{\varphi\}$, and since $c[\varphi]$ will always satisfy $\varphi$, no matter what information $c$ contains, this conjunction imposes no requirements on its context at all. In particular, the satisfaction theory correctly predicts that (10) doesn't presuppose that Fred is married.

Now compare the following sentences:

(11)    a.    It is $\left\{\begin{array}{l}\text{posssible}\\\text{not true}\end{array}\right\}$ that Fred's wife hates sonnets and that his manager does so, too.

        b.    It is $\left\{\begin{array}{l}\text{posssible}\\\text{not true}\end{array}\right\}$ that Fred hates sonnets and that his wife does so, too.

Intuitively, we would infer from both sentences that Fred has a wife, but the satisfaction theory only predicts this for (11a). Since in each case the highest operator is a hole to presuppositions, (11a) and (11b) are predicted to have the same presuppositions as (8a) and (9a), respectively: (11a) presupposes that Fred is married, but (11b) merely presupposes that Fred has a wife *provided that* he hates sonnets, and in this case it is not at all clear how the stronger inference might be obtained. So in its pristine form the satisfaction theory cannot explain why we would normally infer from (11b) that Fred has a wife. Since this problem arises because a presupposition is weakened by a

condition that is not intuitively observable, I will call this the 'proviso problem'.

The proviso problem arises because what is intuitively speaking presupposed is often more (or stronger) than what is required to guarantee local satisfaction of a given presupposition. The problem isn't restricted to conjunctions but also turns up with conditionals, which isn't so remarkable since conjunctions and conditionals share the same definedness conditions. Thus, the theory predicts that (12a) presupposes that Fred has a wife, while (12b) merely presupposes that he has a wife if he hates sonnets:

(12)    a.    If Fred's wife hates sonnets then his manager does so, too.

        b.    If Fred hates sonnets then his wife does so, too.

Again, although intuitively both sentences imply that Fred has a wife, the satisfaction theory straightforwardly predicts this for (12a) but not for (12b), which is predicted to presuppose merely that the conditional in (9b) is true.

In short, it turns out to be a characteristic of all versions of the satisfaction theory that for a certain range of constructions they deliver predictions that are too weak. This proviso problem arises, in particular, when a presupposition is triggered in the consequent of a conditional or in the second member of a conjunction that occurs in a non-entailed position, for example, in the scope of a negation or a modal operator. The problem isn't restricted to these environments, however. The local context in which a presupposition $\chi$ arises is given by an expression of the form $c[\varphi_1][\varphi_2] \dots [\varphi_n]$, $n \geq 0$, where c is the global context. The general proviso problem is that the satisfaction theory predicts that in this situation only a presupposition of the form $(\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n) \to \chi$ will survive, whereas the intuitive inference will generally be that $\chi$ is true. Thus for the following example, which is of the form $\varphi \to (\psi\{\chi\} \wedge \psi'\{\chi'\})$,

(13)    a.    If Fred writes one more sonnet, his wife will be happy and his manager will be sore.

        b.    Fred writes one more sonnet $\to$ Fred is married

        c.    (Fred writes one more sonnet $\wedge$ Fred's wife is happy) $\to$ Fred has a manager

the satisfaction theory predicts that (13a) presupposes (13b) and (13c) rather than 'Fred is married' and 'Fred has a manager', respectively.

It should have become clear in the foregoing that the proviso problem isn't a problem about conditionals — it is *also* a problem about conditionals. However, in the following, conditionals will play an important role, because if we want to keep the complexity of our data at a minimum, samples of the form $\varphi \to \psi$ are preferable to, say, ones of the form $\neg(\varphi \wedge \psi)$ or $\varphi \vee (\psi \wedge \chi)$. In

a sense it is unfortunate that this should be so, because the semantics of conditionals is a very moot issue indeed. It is generally acknowledged that the natural-language conditional cannot be identified with the material implication, as we have done so far and will do for the time being, and a number of semantic theories are available designed, for example, to avoid the notorious 'paradoxes' of material implication. But although I agree that the material implication is at best a very rough approximation to the meaning of the natural language conditional, I don't think that the problems to be discussed below are alleviated by adopting a more sophisticated semantics of *if . . . then*. To illustrate this point, let us consider the possibility of incorporating into our dynamic semantics the theory of conditionals developed by Stalnaker and Lewis, which surely ranks among the more respectable accounts (see Stalnaker 1968, 1975, Lewis 1973). Formulated in truth-conditional terms, this analysis says that 'if $\varphi$ then $\psi$' is true in the current world iff $\psi$ is true in all those $\varphi$-worlds that are maximally similar to the current world. Adopting a proposal by Heim (1992), this translates into the context-change format as follows, where '>' represents the Stalnaker/Lewis conditional:

(14)     $c[\varphi > \psi] = \{w \in c \mid sim_w(c[\varphi])$ satisfies $\psi\}$, where $sim_w(c[\varphi])$ is the set
         of $w' \in c[\varphi]$ such that $w'$ is maximally similar to $w$

This is defined iff $c[\varphi]$ is defined and for all $w \in c$, $sim_w(c[\varphi])[\psi]$ is defined. So, assuming that $c[\varphi]$ is defined, $c[\varphi > \psi\{\chi\}]$ is defined iff for all $w \in c$, $sim_w(c[\varphi])$ satisfies $\chi$, and this holds iff c satisfies $\varphi > \chi$. That is to say, if we adopt '>' instead of '$\rightarrow$' as the formal counterpart to the natural language conditional, the prediction that $\varphi \rightarrow \psi\{\chi\}$ presupposes $\varphi \rightarrow \chi$ gives way to the parallel prediction that $\varphi > \psi\{\chi\}$ presupposes $\varphi > \chi$. Intuitively, this is still not right. However, while in the former case the presupposition predicted is weaker than the inference that is intuitively perceived, in the latter case the predicted presupposition is neither stronger nor weaker than the observed inference.

   Note that if we decide to represent the natural language conditional by '>' instead of '$\rightarrow$', the theory's predictions about conditionals and conjunctions will diverge: $\varphi > \psi\{\chi\}$ presupposes $\varphi > \chi$ but of course $\neg(\varphi \wedge \psi\{\chi\})$ still presupposes $\varphi \rightarrow \chi$. So unless the interpretation of the conjunction is changed, too, the proviso problem will now appear in two forms: some of the presuppositions that the theory predicts will be too weak, while others will be logically unrelated to what is observed.

   For the time being, I will stick to the dynamic semantics defined in (1), and construe conditionals in terms of material implication. The main reason for doing so is that the proposal by Karttunen and Peters which is discussed in the next section starts out from the premise that the conditionalized presuppositions predicted by the satisfaction theory are truth-functional. In

§ 3.4 an argument is discussed that requires them to be *non*-truth-functional, and then the Stalnaker/Lewis analysis will be taken up again.

It may perhaps be less than obvious why the proviso problem is a problem at all. For intuitively speaking it is perfectly clear how examples like (11b) and (12b) should be analysed: since the presupposition that Fred has a wife is not satisfied in these cases, *that presupposition* (i.e. the presupposition as triggered by the possessive NP) should be accommodated. However, for a satisfaction theorist this isn't an appealing line to take, to say the least. The beauty of the satisfaction theory is that the facts about presupposition projection automatically fall out of a theory of context change and some general principles governing accommodation.[4] According to the satisfaction theorist, there is no need for a *theory* of presupposition projection; all he needs, or so he hopes, is a list of presupposition *inducers*, which records that definite NPs trigger existential presuppositions, that factive verbs presuppose their complements, and so on. Once he has such a list, he expects his dynamic semantics to tell him what requirements any given sentence imposes on a context in which it is uttered. But seen from this perspective, if he is prepared to stipulate that under certain circumstances a stronger proposition must be accommodated than is required by the rules of context change, the satisfaction theorist can dismiss not only the proviso problem, but one of his main premises, to boot.

This point is important enough to be laboured a bit. As we have seen, the satisfaction theory predicts that $\varphi \to \psi\{\chi\}$ presupposes that $\varphi \to \chi$. In fact, this prediction follows automatically from the dynamic semantics given in the previous section. Suppose now that we would rule that, if c doesn't satisfy $\varphi \to \chi$, then $\chi$ must be accommodated. It would be easy, in fact trivial, to postulate this. But if we adopt this measure, there ceases to be a connection between the presuppositions predicted by our dynamic semantics, on the one hand, and what the theory says about accommodation, on the other. So the point is not that a postulate to this effect wouldn't solve the problem, but that the problem would be solved at the expense of the theory's austere elegance.

There are examples which have been argued to show that the conditionalized presuppositions predicted by the satisfaction theory are correct. The following are cases in point:[5]

(15)    a.    If Fred is a scuba diver, then he will bring his wet suit.

        b.    Fred is a scuba diver → Fred has a wet suit

---

[4]  I am prepared to grant that these principles aren't restricted to presuppositional phenomena (cf. Lewis 1979), and can therefore be motivated on independent grounds.
[5]  I owe these examples to Manfred Krifka. See Beaver (1992, 1993), who alleges that these are counterexamples against van der Sandt's theory of presupposition projection.

(16)    a.    It is possible that Fred has diabetes and must eat his regular
              diet.

        b.    Fred has diabetes → Fred must eat his regular diet

It seems possible to read (15a) as implying that Fred has a wet suit, no matter
what. But it is certainly more natural to merely infer from (15a) that (15b) is
true, which is what the satisfaction theory predicts.[6] And for (16a), too, it
seems that the satisfaction theory produces an adequate prediction.
Therefore, these examples present no problems to the theory, but it should
be noted that the support that they offer is more apparent than real. To see
why, consider the following variation on (15):

(17)    a.    If Fred is a scuba diver and wants to impress his girlfriend,
              then he will bring his wet suit.
        b.    (Fred is a scuba diver ∧ Fred wants to impress his girlfriend)
              → Fred has a wet suit

(17a) is the result of strengthening the antecedent of (15a), and intuitively
this has no consequences for the presuppositions that this sentence gives rise
to: if the favoured presupposition of (15a) is (15b), then the same
presupposition should be associated with (17a). However, in this case the
satisfaction theory cannot offer more than (17b), and I find it difficult to
understand why a theory that makes correct predictions about (15a) should
fail to make sufficiently strong predictions about (17a), unless of course the
correct predictions are simply due to a fortuitous coincidence.

   The proviso problem isn't restricted to the satisfaction theory: it also arises
in Karttunen and Peters's (1979) system and in various presuppositional
logics. If we construe, say, strong Kleene logic as a presupposition theory,[7]
we find again that $\varphi \to \psi\{\chi\}$ presupposes $\varphi \to \chi$. Predictably, therefore, the
problem hasn't gone unnoticed, but although the proviso problem has often
been mentioned in the literature, I know of only two attempts at solving it.[8]
One proposed solution, first suggested by Karttunen and Peters, will be
discussed in the remainder of this section and in § 3.3; the second proposal is
the subject of § 3.4.

   Karttunen and Peters maintain that there is really nothing wrong with the
predictions that the satisfaction theory makes: it is *correct* that $\varphi \to \psi\{\chi\}$

---

[6]   This particular prediction is pre-empted by the fact that (15b) is already entailed by (15a),
anyway, but that doesn't hold for the next example.
[7]   Quite apart from the problem noted here, this would be a rather unfortunate decision. See
Karttunen and Peters (1979) and Krahmer (1993) for discussion.
[8]   See Karttunen and Peters (1979), Gazdar (1979), Soames (1979, 1982), Heim (1983), van der
Sandt (1988), Krahmer (1993), and Beaver (1993).

merely presupposes that $\varphi \to \chi$ is true.[9] On the other hand, Karttunen and Peters don't want to deny that under normal circumstances $\varphi \to \psi\{\chi\}$ would license a stronger inference, namely that $\chi$ is true. The moral they draw from this is that there must be some train of reasoning which on the basis of, inter alia, the conditional presupposition $\varphi \to \chi$ yields the conclusion that $\chi$ is true. This inference is not itself of a presuppositional nature, the 'real' presupposition $\varphi \to \chi$ being just one of the premises on which it is based. In the following I shall first criticize the general outlines of Karttunen and Peters's account, and then turn to two arguments that have been proposed in order to explain the alleged strengthening of conditional presuppositions.

Regardless what the details of the strengthening process look like, Karttunen and Peters's proposed solution to the proviso problem is a priori implausible. They assume a systematic difference between, for example, sentences of the form $\varphi\{\chi\} \to \psi$ or $\neg(\varphi\{\chi\} \wedge \psi)$ on the one hand and $\varphi \to \psi\{\chi\}$ or $\neg(\varphi \wedge \psi\{\chi\})$ on the other, and this difference is simply not there, intuitively speaking: introspectively each of these forms licenses the inference that $\chi$ is true (still assuming that $\varphi$ and $\chi$ are logically independent), and there is simply no pre-theoretical evidence in favour of the assumption that this inference has different sources depending on whether the presupposition is triggered in the first or in the second half of the sentence. Karttunen and Peters's argument is therefore quite clearly an attempt to reason away the unwelcome side effects of the satisfaction theory.

### 3.3 The argument from truth-functionality

Let us consider how Karttunen and Peters's proposal might be fleshed out. I shall first outline the argument that Karttunen and Peters themselves propose, and then present a number objections against it. In the next section, I will consider another argument, and attempt to show that most of these objections apply to it as well. Karttunen and Peters (1979) propose a quasi-Gricean argument for deriving $\chi$ from the presupposition that $\varphi \to \chi$. The general pattern of this argument from truth-functionality, as I call it, is the following (see also Soames 1979, 1982):

(18)    *The argument from truth-functionality*

   a.    The speaker has uttered a sentence of the form $\varphi \to \psi\{\chi\}$ and therefore he has *said* that $\varphi \to \psi$ and *presupposed* that $\varphi \to \chi$.

[9]   As I noted above, unlike Karttunen's 1974 proposal, the theory that Karttunen and Peters present in their 1979 article is not a satisfaction theory as I understand that notion. Empirically, however, their theory runs afoul of the same trouble as Karttunen's earlier theory. So originally the following argument was not intended to save the satisfaction theory but rather Karttunen and Peters's.

  b.   It may be assumed that the speaker's grounds for presup-
       posing φ → χ are of a truth-functional nature — i.e. he knows
       either that φ is false or that χ is true.

  c.   It may be assumed that the speaker doesn't know if φ is true or
       false (this is conversationally implicated by his saying φ → ψ),
       and therefore it may be assumed that he doesn't know that φ
       is false.

  d.   Therefore, the speaker knows that c is true.

This version only applies to conditional sentences, but it is easily extended to
other cases. The two essential clauses in this schema are (18b) and (18c).
Note that the latter is dependent upon a conversational implicature, but not
necessarily so: the argument would still apply if we had other grounds for
assuming that the speaker doesn't know that φ is false (the relevance of this
point will become clear later on).

   To illustrate how this argument is supposed to work, let us consider the
following example, which is discussed by Soames (1982):

(19)   a.   If the problem was difficult, then Morton isn't the one who
            solved it.

       b.   the problem was difficult → someone has solved the problem

       c.   Someone has solved the problem.

The consequent of (19a) contains an *it*-cleft which triggers the presup-
position that (19c) holds, and therefore the satisfaction theory predicts that
the sentence as a whole presupposes that (19b) is true. Since in most
contexts, a problem's being difficult will not count as a reason for supposing
that someone has solved it, the speaker is presumably presupposing (19b) on
truth-functional grounds. But by asserting (19a) he has conversationally
implicated that he doesn't know if the problem was difficult, and we must
therefore assume that he doesn't take (19b) to be true by falsity of
antecedent. Therefore, it must be the case that he takes (19c) to be true.

   I have four objections against this argument. First, the argument only goes
through if we may plausibly assume that the speaker has truth-functional
grounds for presupposing a conditional. Unfortunately, however, the
inference that (18) seeks to account for will in general go through even when
the speaker is likely to have *non*-truth-functional grounds for making this
presupposition. To illustrate this, we need only slightly change Soames's
example:

(20)   a.   If the problem was easy, then Morton isn't the one who solved
            it.

       b.   the problem was easy → someone has solved the problem

All I have done in (20a) is replace the adjective *difficult* with its antonym, but the strong inference still goes through: (20a) implies that someone solved the problem just as much as (19a) does. However, the conditional presupposition that the satisfaction theory predicts is (20b), and in this case it is perfectly plausible that the speaker doesn't have truth-functional grounds for making this presupposition: if a problem is easy the degree of probability that it has been solved is relatively high. And since, furthermore, non-truth-functional interpretations are generally preferred to truth-functional ones, the second step in the inference schema, (18b), surely doesn't go through, and thus the truth-functionality argument is inapplicable to (20a).

My second objection against the argument from truth-functionality concerns the assumption that the speaker does not presuppose (19b) because he takes the antecedent to be false. In the examples that we have considered thus far this premise is derived as a conversational implicature: since the speaker has uttered a sentence of the form $\varphi \rightarrow \psi$ his audience are entitled to assume that he doesn't know that $\varphi$ is false (or true). As this is a conversational implicature it can be overwritten: the speaker might simply say that he knows the antecedent to be false, for example. But if he chooses to do so, the inference schema in (18) no longer holds, and we would expect the inference to be retracted as well. That is not what happens, however.

(21)    If the problem was difficult, then Morton isn't the one who solved it. But as a matter of fact the problem wasn't difficult at all.

In (21), the speaker explicitly asserts that the problem wasn't difficult, thus overwriting the conversational implicature that his first utterance gives rise to. Consequently, condition (18c) is no longer satisfied, and the inference should not go through anymore. But of course we would still infer from the discourse in (21) that the speaker assumes that someone solved the problem. Again, this suggests rather strongly that nothing like the truth-functionality argument is involved in this inference.

The third objection concerns the same clause in the argument, i.e. (18c). This clause makes the desired inference contingent upon a conversational implicature (at least in the cases we have discussed so far), which leads us to expect that the inference is cancellable. We saw in (21) how the conversational implicature licensed by the first member of a conditional (in this case, that the speaker doesn't know if the problem was difficult or not) can be cancelled. Now compare that discourse with the following:

(22)    ?If the problem was difficult, then Morton isn't the one who solved it. But as a matter of fact the problem wasn't solved at all.

Or compare the following two discourses:

(23)    If there are piranhas in the Rhine, then Fred's wife should know about it.

    a.    But of course there are no piranhas in the Rhine.

    b.    ?But of course Fred isn't married.

It seems to me that while it is perfectly okay to cancel an ordinary conversational implicature, it is impossible to cancel the inference which, according to the satisfaction theory, is based upon such an implicature.

    The observation which the fourth and final objection turns upon is a straightforward one. The argument from truth-functionality is meant to strengthen conditional presuppositions which arise, or are supposed to arise, in an indirect way, as they are pieced together out of material contributed by a presuppositional expression as well as its carrier sentence (for example, the antecedent of a conditional). But actually it shouldn't make a difference *how* this presupposition arises. In particular, the argument should also apply if the presupposition were triggered directly, for example, if a conditional is embedded within the scope of a factive predicate, as in the following example:

(24)    Barney knows that if the problem was difficult, then someone solved it.

Here the presupposition that (19b) is true is triggered directly, and intuitively (24) does indeed presuppose that (19b) is true, *and nothing more.* In fact it would be quite remarkable if (24) would ever give rise to the inference that someone solved the problem, except of course in contexts in which it is given that the conditional's antecedent is true. According to the satisfaction theory, however, (24) parallels (19a) and it should therefore be possible to find contexts in which it implies that someone solved the problem. In fact, it should be sufficient that a speaker who utters (24) doesn't know if the problem is difficult, for then the two essential conditions in the truth-functionality argument, viz. (18b) and (18c), would be satisfied. This prediction doesn't tally with our intuitions, however: it is obvious that even in such a context, we wouldn't infer from (24) that someone solved the problem.

    Or consider the following discourse:

(25)    I don't know if the problem was difficult or not, but I do find it surprising that if the problem was difficult, then someone has solved it.

A speaker who utters (25) presupposes that (19b) is true and asserts that he doesn't know if the problem is difficult or not. But we wouldn't normally infer from his utterance that he assumes that someone has solved the problem. So, once more, the truth-functionality argument makes the wrong predictions.

I believe that the contrast between (19a) and (24) is of considerable interest, because it will present a problem to practically any theory which claims that the inference from (19a) to (19c) is mediated by the presupposition in (19b). As the following examples show, this contrast is produced not only by *it*-clefts but by other presuppositional expressions as well:

(26)    a.    If Barney is sensible, then he will stop smoking.

        b.    Fred knows that if Barney is sensible, then he is a habitual smoker.

        c.    Barney is sensible → Barney is a habitual smoker

        d.    Barney is a habitual smoker.

(27)    a.    If Julius had canard à l'orange, then what his wife ate was potato chips.

        b.    Fred knows that if Julius had canard à l'orange, then his wife ate something (too).

        c.    Julius had canard à l'orange → Julius's wife ate something

        d.    Julius's wife ate something.

(28)    a.    If Betty is married, then Wilma's husband should be more discreet.

        b.    Fred knows that if Betty is married, then Wilma has a husband (too).

        c.    Betty is married → Wilma has a husband

        d.    Wilma has a husband.

Intuitively, we would infer from an utterance of any of the (a) sentences in (26)-(28) that the speaker takes the corresponding (d) sentence to be true as well; in (26a) this presupposition is triggered by the aspectual verb, in (27a) by the *wh*-cleft, and in (28a) by the definite NP *Wilma's husband*. In contrast, these inferences would not normally follow from the (b) sentences. However, the satisfaction theory predicts that, in each of these cases, the (a) and (b) sentences have the same presupposition, viz. (c), and that in the first case it is strenghtened to (d). Why is this presupposition strengthened in the former case but not in the latter? Clearly, this is bound to be an embarrassing question not only to theories that adopt the argument from truth-functionality, but to virtually any attempt at dealing with the proviso problem.

The problem exemplified by (26)-(28) is unlikely to go away if we adopt an alternative analysis of conditionals. Thus far we have construed *if . . . then*

sentences in terms of material implication, and it is well known that this construal is a problematic one. However, as we have seen in § 3.2, the proviso problem will remain even if we adopt the Stalnaker/Lewis analysis of conditionals, for example. What this means is that if we adopt this construal, the satisfaction theory still predicts that the (a) and (b) sentences have the same presupposition, the only difference being that the conditional is no longer construed as material implication but as a relation between possible states of affairs. It would be premature to conclude from this observation that there is no analysis of conditionals that will solve the problem, but it does suggest that this line of defence isn't very promising, either.

Another way of approaching the problem would be to deny that the conditionals involved in the (a) and (b) sentences in (26)-(28) are of the same type, and to argue on this basis that these sentence pairs don't give rise to the same presuppositions. There are at least two problems with this idea. First, it will have to be shown on independent grounds that in the (a) sentences *if . . . then* is construed differently than in the corresponding (b) sentences, and it is by no means obvious how that can be done. Secondly, it is clear that such an argument wouldn't work for the following minimal pairs:

(29)    a.    Maybe Fred knows that if Barney is sensible, then he will stop smoking.

        b.    Maybe Fred knows that if Barney is sensible, then he is a habitual smoker.

(30)    a.    Maybe Fred knows that if Betty is married, then Wilma's husband should be more discreet.

        b.    Maybe Fred knows that if Betty is married, then Wilma has a husband (too).

Intuitively, while (29a) presupposes that Barney is a habitual smoker, (29b) doesn't seem to have this presupposition, and analogously, (30a) but not (30b) presupposes that Wilma has a husband. But of course there is no reason whatsoever for believing that the conditionals in the (a) sentences are construed differently from those in the (b) sentences.

## 3.4  The argument from improbability

I believe that the objections that I have mustered in the foregoing warrant the conclusion that the argument from truth-functionality is not just wrong but entirely on the wrong track: the problems it faces are so serious that it calls for replacement rather than repair. One possible replacement might be the following argument, which is based on the same assumptions as

Karttunen and Peters's: it takes as its point of departure the idea that the predictions of the satisfaction theory are correct, but occasionally require strengthening. The argument goes as follows:[10]

(31)  *The argument from improbability*

      a.    The speaker is presupposing that $\varphi \to \chi$.

      b.    It is more plausible (or less controversial) to assume that $\chi$ than to assume that $\varphi \to \chi$.

      c.    Therefore, the speaker is probably assuming that $\chi$.

This is similar to the first argument in that both assume that the conditional presuppositions which cause the proviso problem are deficient in some way, but in (31) Karttunen and Peters's original idea that these conditionals are purely truth-functional has watered down to the assumption that they are comparatively unlikely, or more controversial than their consequents.

Prima facie, the argument from improbability seems to be much more plausible than the argument from truth-functionality, but on closer inspection it turns out to be equally inadequate. To begin with, the argument requires that $\chi$ be more plausible than $\varphi \to \chi$, but it is obviously impossible for a proposition to be more plausible than a proposition which it entails. Therefore, this argument will not work unless we give up our construal of natural language conditionals in terms of the material implication. For example, if we adopted '>' as defined in (14) instead of '$\to$', 'if $\varphi$ then $\psi\{\chi\}$' would presuppose $\varphi > \chi$, which is neither weaker nor stronger than $\chi$. Then it might be possible to claim that $\varphi > \chi$ is more plausible than $\chi$, and the argument outlined in (31) might go through.

However, although '>' surely comes much closer to the way conditionals are construed in everyday practice, adopting it instead of the material implication in this case causes more problems than it solves.[11] First, as we have seen in § 3.2, even if we decide to construe conditionals in terms of '>', the satisfaction theory will still give rise to presuppositions of the form $\varphi \to \chi$ where we would like to have $\chi$. This is because conjunctions are not (and, evidently, shouldn't be) affected by the decision to re-interpret the conditional. For example, although $\varphi > \psi\{\chi\}$ presupposes $\varphi > \chi$, $\neg(\varphi \wedge \psi\{\chi\})$ still presupposes $\varphi \to \psi$. So the argument from improbability will solve only part of the problem, at best. I say 'at best' because it is easily seen that the argument will not work for conditionals, either. One of the so-called 'paradoxes' of the material implication is that it is weaker than its

---

[10] The provenance of the argument from improbability is unclear to me. I have the impression that many people have taken it for granted that something like it would work, but the only printed version that I know of is Beaver's (1992, 1993), which will be discussed below.

[11] The problem to be discussed in the following extends to non-extensional contexts in general.

consequent: $\psi$ entails $\varphi \to \psi$. Clearly, the semantics of *if . . . then* shouldn't imply that a conditional and its consequent are logically related, and one of the advantages of the Stalnaker/Lewis theory is precisely that it manages to avoid this conclusion. However, if we introduce a conditional with this property to replace the material implication, the proviso problem is actually exacerbated. As long as the satisfaction theorist can say that his dynamic semantics produces presuppositions that are just sometimes too weak, his predicament at least makes sense. But, ironically, the plausibility of his position diminishes as soon as he adopts a semantics of conditionals that is more adequate than what the material implication has to offer. In general, if a sentence of the form $\varphi > \psi\{\chi\}$ has just been uttered in a context in which $\varphi > \chi$ is not yet given, the hearer's accepting $\chi$ doesn't contribute towards the satisfaction of $\chi$ in its local context. Regardless whether he accepts $\chi$ or not, he must *also* accommodate $\varphi > \chi$ (or something stronger). But then what is the point of inferring $\chi$ on the basis of $\varphi > \chi$, even if the former is more plausible than the latter? Or, to put it the other way round, the satisfaction theory implies that even if $\chi$ is contextually given, the presupposition in $\varphi > \psi\{\chi\}$ isn't automatically satisfied — a consequence which strikes me as paradoxical if not absurd.

Even if we set these difficulties aside, there are further problems that defenders of the improbability argument will have to face. Since it is based on comparative plausibility judgments, the argument implies that there is a parallel between our intuitions about the presuppositions of sentences of the form, for example, $\varphi \to \psi\{\chi\}$ on the one hand, and our intuitions about the relative plausibility of $\chi$ and $\varphi \to \chi$ on the other.[12] As I observed in § 3.3, (32a) strongly suggests that (32c) is true, and the argument from improbability seeks to explain this fact on the premise that (32c) is more plausible than (32b), which according to the satisfaction theory is presupposed by (32a).

(32)    a.    If the problem was easy, then Morton isn't the one who solved it. (= (20a))

        b.    the problem was easy $\to$ someone has solved the problem (= (20b))

        c.    Someone has solved the problem. (= (19c))

I am not going to quibble about the relative probabilities of (32b) and (32c) (although, off the record, I don't see that the latter is more plausible than the former). My point is the following. On the one hand it is obvious that (32a)

---

[12]  I reintroduce the arrow here because I have been using it for most of the time, and as I have just argued, '$\to$' and '$>$' are equally problematic in the present context. The following remarks don't depend on any particular way of construing conditionals.

implies (32c), in a fairly strong sense of 'imply'. On the other hand it is anything but obvious that (32c) is more plausible than (32b). However, if the argument from improbability is correct, then the two preceding statements cannot be true together, because judging that (32c) is more plausible than (32b) is tantamount to judging that (32a) implies (32c). But since these two statements are, evidently, not the same thing, the argument must be wrong. Nor is this problem restricted to a single example:

(33)    a.    If Mwamba is a monarchy, then Mzamba is a monarchy (too).

           b.    Mwamba is a monarchy → Mzamba is a sovereign state

           c.    Mzamba is a sovereign state.

(34)    a.    If Mwamba is a monarchy, then the king of Mzamba will support it.

           b.    Mwamba is a monarchy → Mzamba has a king

           c.    Mzamba has a king.

(35)    a.    If Mwamba is a monarchy, then the king of Mzamba will continue to support it.

           b.    Mwamba is a monarchy → until now, the king of Mzamba has supported Mwamba

           c.    Until now, the king of Mzamba has supported Mwamba.

The argument from improbability requires that the strong intuition that the (a) sentences in (33)-(35) imply the corresponding (c) sentences should be matched by an equally strong intuition that the (c) sentences are more plausible than the corresponding (b) sentences. But there is no such match.

One of my objections against the truth-functionality argument was that it makes the strengthening of conditional presuppositions contingent upon conversational implicatures, and thus leads us to expect, wrongly, that these inferences are cancellable. Although the argument from improbability doesn't rely upon conversational implicatures, the foregoing observations suggest that a similar objection applies in this case, too. The reason is that this argument crucially relies upon judgements of relative probability, which aren't always clear cut. But if an inference is based upon uncertain premises, we shouldn't expect it to be particularly solid. And as we have seen, this expectation is simply incorrect. There are differences of a presuppositional nature between sentences of the form $\varphi \to \psi\{\chi\}$ and $\varphi\{\chi\} \to \psi$,[13] but it is not that *if* we conclude that $\chi$ from either of these, the inference is more

---

[13] In particular, in $\varphi \to \psi\{\chi\}$, but not in $\varphi\{\chi\} \to \psi$, the presupposition that $\chi$ may be 'blocked' in the antecedent of the conditional.

easily cancelled, or less secure, in the former case than in the latter. With respect to the examples in (32)-(35), I think it is clear that most hearers would confidently judge that the (a) sentences imply the corresponding (c) sentences. But they would not nearly be as confident in their judgements about the comparative probabilities of the (b) and (c) sentences. But then it seems rather unlikely that the former judgements should be based, inter alia, upon the latter.

Of course, I don't want to argue that projection phenomena are completely impervious to (certain forms of) uncertain knowledge; an example like (36) suffices to prove that this isn't the case (cf. also the discussion of bridging in § 2.4.3).

(36)      If Fred is a scuba diver, then he will bring his wet suit. (= (15a))

It is *apparent* that, if we don't infer from this that Fred has a wet suit, then it is because we assume that scuba divers often if not always possess wet suits. But if the inference from (32a) to (32c) would likewise be based upon some piece of world knowledge, we should expect this to be equally apparent. Which it isn't.

Finally, the improbability argument applies to (37a), and delivers the patently false prediction that we would infer from this sentence that someone solved the problem.

(37)      a.      Barney knows that if the problem was difficult, then someone
                  solved it. (= (24))

          b.      If the problem was difficult, then Morton isn't the one who
                  solved it. (= (19a))

This counterexample is even more problematic for the present argument because we don't have to make any special provisions about the context in which (37a) is uttered: if the argument from improbability applies to (37b), it should apply to (37a) as well, because according to the satisfaction theory the relevant presupposition is the same in both cases. Conclusion: although prima facie it is less artificial than the argument from truth-functionality, the improbability argument doesn't fare any better.

To conclude my discussion of the proviso problem, I want to briefly consider a version of the argument from improbability proposed by Beaver (1992, 1993). Thus far I have construed this argument as based upon a plausibility measure of propositions: it is because the hearer believes that $\varphi \to \chi$ is less plausible than $\chi$ that he is prepared to infer that $\chi$. Beaver claims, however, that the relevant plausibility judgements don't concern the hearer's beliefs about the world, but rather about what he takes to be the common ground between the speaker and himself. The idea is that in general it is not fully determinate what is in the common ground and what is not (at any given point in the discourse). Rather, each of the interlocutors may

entertain several hypotheses about the common ground, some of which he may deem to be more plausible than others. An interesting consequence of this proposal is that it opens a new perspective on one of the problems that I discussed in the foregoing. I observed that it is difficult to see how $\chi$ can be more plausible than the weaker $\varphi \rightarrow \chi$, as the argument from improbability requires, but Beaver would say that this difficulty only arises because I have been asking the wrong question. What I should have asked instead, according to Beaver, is whether the hearer finds it more likely that the *common ground* satisfies $\chi$ or just $\varphi \rightarrow \chi$. And the hearer's answer to this question might very well be that the former possibility is more likely than the latter.

Although Beaver's proposal offers a solution to one of the problems that have come up in the foregoing, I don't see that it improves matters in others than this respect. Beaver shows how the satisfaction theorist can consistently adopt the argument from improbability without having to give up on the standard dynamic semantics of § 3.1. However, as I have pointed out in this section, it is clear that at least for conditionals this semantics is inadequate, and it still remains to be seen how a more sophisticated semantics of conditionals can be incorporated into the satisfaction framework. Secondly, Beaver's version of the improbability argument does nothing to alleviate the problems illustrated by (32)-(35). Indeed, it seems to me that if, instead of asking whether the (c) propositions are more plausible than the (b) propositions, we now ask which ones are more likely to be part of the common ground, it becomes even more questionable that there will be a majority vote for (c), as the argument from improbability requires. And finally, Beaver's version of the argument of improbability doesn't account for the difference between, for example, (37a) and (37b).

Summarizing the results of this section and the previous two, I have argued against the idea that the proviso problem can be solved along the lines first suggested by Karttunen and Peters. I have objected against the very idea of deriving the required inferences in two steps, and have criticized two arguments that have been proposed to account for the alleged strengthening of conditional presuppositions. My main conclusions are: that the conditional presuppositions which the satisfaction theory predicts are mere artefacts of that theory; that the notion that the observed inferences are to be derived in a two-step process is ill conceived; and that there is at present no account that explains how this strengthening might be accomplished.

## 3.5  Further problems

The main problem with the satisfaction theory is that certain of the presuppositions that it predicts are conditionalized whereas intuitively they

shouldn't be. But occasionally, the theory simply fails to generate any pre-suppositions at all while intuitively we would like to have some. This problem arises because the theory predicts that sentences of the form $\varphi \rightarrow \psi\{\chi\}$ or $\varphi \wedge \psi\{\chi\}$ will not presuppose anything if $\chi$ already follows from $\varphi$. This gives the right predictions in many cases, but it is not always correct, as the following examples show:

(38)     a.     If all countries have presidents, then the president of France
                  probably regards himself as their cultural leader.

         b.     It is possible that all countries have presidents and that the
                  president of France regards himself as their cultural leader.

Example (38a) is taken from Soames (1982: 539), who argues that the satisfaction theory gives the correct predictions in this case, and if Soames were right about this, the same would hold for (38b). Soames's reasoning is defective though. His argument is based upon the observation that (38a) 'is *neutral* regarding whether or not the speaker assumes [that France has a president], and no utterance presupposition is heard,' (ibid.) and he concludes from this that a theory which doesn't predict a presupposition in this case has the facts right. This conclusion isn't warranted, however, because only half of the observation on which it is based is correct. Soames is right in saying that, taken by itself, (38a) remains neutral with respect to the presupposition that France has a president, but this doesn't mean that it never will have this presupposition, as the satisfaction theory predicts. Rather, (38a) can be read either as presupposing or as not presupposing that France has a president, and the satisfaction theory only accounts for the latter possibility.

   That both sentences in (38) have a reading on which it is presupposed that France has a president can be seen from the fact that this individual may be taken up by an anaphoric pronoun. If either (38a) or (38b) were followed by an utterance of,

(39)     He is such a pompous ass.

it would be perfectly clear who was being referred to. But the examples in (38) will not license this anaphoric link unless they are construed as presupposing that France has a president. It is this strong reading that the satisfaction theory fails to account for.

   It may be objected that the problem presented by the examples in (38) is only an apparent one, and will dissolve once we allow for the possibility that the definite NP *the president of France* has a referential interpretation. For if this NP is construed referentially, it outscopes the conditional to begin with, and the seemingly problematic reading is accounted for in a straightforward

manner. I will not discuss this suggestion in detail,[14] and merely wish to point out that the problem exemplified by the examples in (38) isn't restricted to definite NPs but can be replicated with all kinds of presupposition inducing expressions and constructions. For example,

(40)    a.    If all the boys failed the exam, then it wasn't only Fred who did so.

         b.    If all the boys left together, then the janitor will not have noticed that Fred left.

If Fred is one of the boys, then (40a) is an obvious truth, but nonetheless this sentence may be construed as implying that Fred failed the exam. However, the presupposition that Fred failed the exam, which is triggered in the consequent of the conditional, is satisfied in its local context, and therefore the satisfaction theory cannot account for this inference. Similarly, if all the boys left together, and still assuming that Fred is one of the boys, then Fred must have left, too. So the presupposition induced by the factive in (40b) is satisfied in its local context, and therefore the satisfaction theory doesn't account for the more obvious interpretation of this sentence, according to which Fred left.

The problem that I have discussed in the foregoing arises because, intuitively, a presupposition may project to the global context although it is locally satisfied. The second problem that I want to bring up is in a sense the mirror image of this: occasionally, a presuppositional expression may be felicitously used although the presupposition it triggers is inconsistent with its local context. The following are cases in point:

(41)    a.    Wilma isn't married. So it wasn't Wilma's husband who shot the burglar.

         b.    If Wilma isn't married, it wasn't Wilma's husband who shot the burglar.

Let $\chi$ = 'Wilma is married' and c an arbitrary context which is consistent with $\chi$ as well as with $\neg\chi$. Operating on this c, the first sentence in (41a) results in $c' = c - c[\chi]$, and thus $c'$ consists of all and only those worlds in c in which Wilma isn't married. We now attempt to increment this context with the second sentence, which is of the form $\neg\varphi\{\chi\}$: $c'[\neg\varphi\{\chi\}] = c' - c'[\varphi\{\chi\}]$, and $c'[\varphi\{\chi\}] = c'[\varphi]$ provided $c'$ satisfies $\chi$. But since $c'$ satisfies $\neg\chi$, $c'[\varphi\{\chi\}]$ is undefined, and so $c'[\neg\varphi\{\chi\}]$ is undefined as well. Therefore, $\chi$ will have to be accommodated, and since the global context $c'$ already satisfies $\neg\chi$, it is clear

---

[14] Though see §§ 1.4 and 5.2 for related discussion.

that local accommodation is called for.[15] That is to say, in order to evaluate $c'$ – $c'[\varphi\{\chi\}]$, we accommodate $\chi$ at the stage at which $c'[\varphi\{\chi\}]$ is computed, so instead of $c' - c'[\varphi\{\chi\}]$, we compute $c' - c'[\chi][\varphi\{\chi\}]$. Now we find that $c'[\chi]$ is empty because $c'$ already satisfies $\neg\chi$, and therefore $c'[\chi][\varphi\{\chi\}]$ is empty, too, and so $c' - c'[\chi][\varphi\{\chi\}] = c'$. This outcome is correct, because the second sentence of (41a) doesn't provide any new information in a context in which the first one is already accepted. Similarly, if we allow for local accommodation as in (41a), (41b) is predicted to be a tautology, which is arguably correct, too.

However, the way these results are obtained is a dubious one, to say the least. The problem is that in both cases, local accommodation results in an empty, i.e. inconsistent, local context, which intuitively speaking doesn't seem to be right. And sure enough, a closer look at the examples in (41) reveals that there are serious problems with this account. Note that, apart from *Wilma's husband*, both examples contain a definite NP that we have ignored so far, viz. *the burglar*. This NP triggers the presupposition that there was a burglar, and accordingly both (41a) and (41b) imply that there was a burglar (again, analogous examples can be constructed with other presuppositional expressions). The satisfaction theory doesn't account for these inferences, however. The local context in which *Wilma's husband* and *the burglar* occur is the same in both cases, and we have just seen that the local context in which the former presupposition is evaluated must be empty. But then this context satisfies any proposition, and in particular that there was a burglar. Therefore, the presupposition that there was a burglar is satisfied in its local context, and doesn't impose any requirements on the global context.

In § 2.4.2, I have shown how the binding theory treats examples like (41a, b) in terms of local accommodation. The foregoing argument demonstrates that this explanation cannot be taken over into the satisfaction framework. As it stands, the satisfaction theory doesn't have a workable account of these cases at all.

According to the satisfaction theory, presuppositions are definedness conditions: unless further possible sources of undefinedness are taken into account, it is necessary as well as sufficient for $c[\varphi]$ to be defined that all presuppositions in $\varphi$ are satisfied in their local contexts. The proviso problem shows, however, that local satisfaction isn't sufficient, and the first problem discussed in this section points in the same direction. The second problem shows that local satisfaction isn't necessary either, for, apparently, what causes the problems with (41) is the requirement that the presupposition that Wilma is married must be satisfied by the local context in which it is

[15] This analysis has been suggested by Heim (1983), Soames (1989), and Beaver (1993).

triggered, which in these cases can be guaranteed only by emptying the local context altogether.

## 3.6 Diagnosis

None of the problems discussed in the foregoing arise in the context of the binding theory. This is a remarkable fact, because the two theories appear to be very similar, and therefore it will be instructive to try and find out why their predictions are so different. The exposition of the satisfaction theory in § 3.1 should have made it clear that there are very close parallels between it and the binding theory. To begin with, the theoretical frameworks in which they are couched are similar, and although I will argue in the next chapter that their similarities should not be given too much weight, there is no denying that DRT and dynamic semantics are related approaches. Both the satisfaction theory and the binding theory construe presupposition in terms of contextual givenness: they both claim that, in principle, a presupposition must be contextually given, and both invoke accommodation as a means to restore givenness when necessary. Furthermore, they impose the same basic restriction on accommodation, viz. that under normal circumstances it must be global. But the import of the two central notions of contextual givenness and accommodation differs between the binding theory and the satisfaction theory.

The binding theory views a presupposition as an *object*, which is expected to be present in some accessible DRS. Accommodation, according to this theory, does not (or not in the first instance) restore definedness; rather it enables an anaphoric link which, in the absence of a suitable antecedent, wasn't possible before. Hence, if no appropriate antecedent object can be found, it is natural to assume that an ersatz must be provided somewhere on the path of DRSs where one was expected. Put otherwise, in view of their object-like nature, it follows naturally that, if a presupposition is to be 'saved' via accommodation, it is the presupposition itself that must accommodated.

This is quite different from the way givenness and accommodation are viewed by the satisfaction theory. This theory construes presuppositional givenness in terms of definedness conditions that are imposed upon the local context in which a presuppositional expression occurs. If a presupposition isn't satisfied it is not because the local context fails to provide an object that was expected to be there, but rather because it doesn't contain the right kind of *information*. In a sense, therefore, an expression like *Fred's wife* triggers different presuppositions according to the two theories: whereas for the binding theory it is that a woman must be given who is married to Fred, for the satisfaction theory it must be contextually given that Fred is married.

So according to the satisfaction theory, what we accommodate is information rather than objects or representations of objects. This difference is not by itself decisive, however. The decisive difference lies in the way a link is established between presupposed and accommodated information. In the binding theory this connection is simple enough: it is the presupposition itself that must be accommodated, which is natural because presuppositions are viewed as objects. However, the satisfaction theory views a presupposition not just as information but as information which is needed *in the local context* in which the presupposition arises. The local context of a presupposition $\chi$ is denoted by an expression of the form $c[\varphi_1][\varphi_2] \dots [\varphi_n]$, $n \geq 0$, and if the *global* context c must be revised so as to let $\chi$ be defined in its *local* context $c[\varphi_1][\varphi_2] \dots [\varphi_n]$, then the theory is forced to predict that it is sufficient if we add $(\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n) \rightarrow \chi$ to c. Thus presuppositional requirements are inevitably weakened because of the way they are transmitted to the global context.

Although this shouldn't be taken too literally, the binding theory pictures presuppositions as agile creatures eager to leave their homes immediately after they have been triggered, in search for suitable antecedents. The satisfaction theory, on the other hand, pictures presuppositions as lethargic beings that keep hanging around in the neighbourhood, content to get local satisfaction. The arguments presented in this chapter refute the latter view, and suggest rather strongly that the former is correct.

CHAPTER 4

# Dynamic semantics

The differences between the binding theory and the satisfaction theory are anything but superficial. We have seen that their views on presupposition diverge considerably, and that, consequently, the two theories yield different predictions. But their differences go even deeper than this; they represent entirely different approaches to meaning. Whereas the binding theory is a version of DRT, the satisfaction theory requires a dynamic semantics; and the differences between these approaches are profound. This is not the standard view. DRT is commonly treated as one among many varieties of dynamic semantics. I maintain that this is misleading, however, and restrict my usage of the term 'dynamic semantics' accordingly.

In this chapter I want to compare DRT and dynamic semantics, and refute the latter. To this end I will concentrate on anaphoric phenomena; presuppositions will not play an important role in this chapter. There are two main reasons for this. First, it is possible to bring out the differences between DRT and dynamic semantics even if we confine our attention to the simplest cases of anaphora. Secondly, although I have argued that the standard dynamic account of presupposition is fatally flawed, I don't want to build my case against dynamic semantics on this argument, because it is not necessary for a dynamic semanticist to adopt the satisfaction theory. It is a natural thing to do, and it would be extremely unnatural to dissociate dynamic semantics from the satisfaction theory — but it is not necessary. However, if one adopts the satisfaction theory, then dynamic semantics is part of the bargain, so any arguments against the latter undermine the former, too.

In the following I first outline the central tenets of dynamic semantics, and present my main objection against the theory (§ 4.1). This objection is perfectly straightforward and, as far as I can see, incontrovertible. Dynamic semantics is an attempt to explain certain pragmatic phenomena in terms of the meanings of the connectives, and this whole approach is plainly wrong. This will be illustrated by means of the dynamic account of conjunction. In the next section I compare dynamic semantics with DRT, and show where the principal differences between these frameworks lie (§ 4.2). In the remainder of the chapter I consider how dynamic semantics might try to explain forward reference (§ 4.3) and certain interactions between anaphora on the one hand, and negation and disjunction on the other (§

4.4), arguing in each case that a plausible dynamic explanation is too much to ask for, because these phenomena, and anaphora in general, call for a pragmatic account.

## 4.1  The theory, and what is wrong with it

In the last chapter we could make do with a propositional version of dynamic semantics, but when we turn our attention to anaphoric phenomena a quantified version is called for. One of the better-known representatives of dynamic semantics is Groenendijk and Stokhof's dynamic predicate logic (DPL), and in the following I will use DPL to exemplify the dynamic approach. This choice is more or less arbitrary; as far as the purposes of this chapter are concerned, I might have chosen any other dynamic theory of quantification instead.

Officially, the syntax of DPL is that of predicate logic, but in the following I will use the semicolon to symbolize conjunction, because I want to emphasize the fact that DPL's dynamic conjunction is quite unlike ordinary conjunction. The interpretation of this language is given in terms of relations between variable assignments, which are to be taken as rudimentary contexts; they replace the sets of worlds that were used in the foregoing. In the following definition, these relations, for which I adopt an infix notation, are symbolized by $[\![.]\!]$. Further, I use $a$, $b$, $c$, ... to represent assignments; I is the interpretation function given by the current model (which is left implicit); $[\![\alpha]\!]_a = I(\alpha)$ if $\alpha$ is an individual constant, and $[\![\alpha]\!]_a = a(\alpha)$ if $\alpha$ is a variable; and '$a[x]b$' is to mean that assignments $a$ and $b$ are identical except, possibly, for the values they assign to $x$.

(1)     *DPL Semantics*

    a.     $a[\![P\alpha_1 \dots \alpha_n]\!]b$ iff $a = b$ and $I(P)([\![\alpha_1]\!]_a, \dots, [\![\alpha_n]\!]_a)$

    b.     $a[\![\neg\varphi]\!]b$ iff $a = b$ and $\neg\exists c: a[\![\varphi]\!]c$

    c.     $a[\![\varphi ; \psi]\!]b$ iff $\exists c: a[\![\varphi]\!]c$ and $c[\![\psi]\!]b$, i.e. $[\![\varphi ; \psi]\!] = [\![\varphi]\!] \circ [\![\psi]\!]$

    d.     $a[\![\exists x\varphi]\!]b$ iff $\exists c: a[x]c$ and $c[\![\varphi]\!]b$

The other connectives and the universal quantifier are defined in terms of '¬', ';', and '∃', as usual. In DPL atomic and negated sentences are tests: they either return their input assignments or fail. Conjunctions and existentially quantified sentences are the only ones that may actually change the contexts on which they operate. The existential quantifier assigns a new value to a variable, and conjunction is construed as relational composition, which is to say that the second conjunct 'applies' after the first; conjunction is therefore non-commutative. Owing to these unorthodox interpretations of '∃' and ';',

the existential quantifier may bind a variable that lies outside its syntactic scope. For example, '$(\exists x\varphi) ; \psi$' and '$\exists x(\varphi ; \psi)$' are equivalent.

In a dynamic theory like DPL, the following contrast is accounted for in terms of the lexical meaning of *and*:

(2)    a.    [A man]$_i$ walks in the park and he$_i$ whistles.

       b.    He$_i$ whistles and [a man]$_i$ walks in the park.

Whereas in the first sentence the pronoun can have the indefinite as its antecedent, this doesn't seem possible in the second sentence. DPL explains this difference as follows. The first sentence is translated as '$(\exists xPx) ; Qx$', which is equivalent to '$\exists x(Px ; Qx)$', and gives an adequate representation of the intuitive meaning of the sentence. In the translation of the second sentence, however, which is of the form '$Qx ; \exists xPx$', the first occurrence of $x$ is free, and thus it is predicted that an anaphoric link is not possible.

I reject this analysis and in particular the notion of conjunction upon which it is based, because in my opinion it is obvious that *and* is not dynamic in the sense that its meaning, or indeed any other aspect of its lexical entry, is relevant to an explanation of why anaphora is possible in (2a) but not in (2b). For once I don't use the adverb 'obvious' to heighten the rhetorical impact of my statement: I really believe that it is obvious that *and* is not dynamic, and that, consequently, dynamic semantics is wrong. I have a strong pre-theoretic intuition that if you must explain the contrast between (2a) and (2b), the lexical entry for *and* is about the least promising place to start, and I have found that many people have the same intuition.[1] Somehow, however, the advocates of dynamic semantics have managed to convince themselves that this intuition is just a mistake. It may be helpful, therefore, to expand upon what I take to be a near-truism. The remainder of this section does just that.

Utterances are processed incrementally: word by word, phrase by phrase, clause by clause, sentence by sentence. Presumably, this strategy is forced upon us by our psycho-biological constitution. More likely than not, it is only thus that we could ever cope with the inexorable stream of verbal input that comes our way. The central tenet of dynamic semantics is that this processing strategy is encoded in the lexical entries of certain words, and especially in the lexical meaning of *and*. Thus formulated, it is evident that the very notion of dynamic semantics is quite implausible. For what could ever be the

---

[1] In December 1997 I posted a query to Linguist List in which I sketched the dynamic treatment of conjunction, and asked my fellow subscribers if they knew of any cross-linguistic evidence which might be taken to confirm this treatment. The outcome was that there doesn't seem to be such evidence. This didn't come as a surprise to me. What I found more interesting was that, judging from the responses I received, the average reader of Linguist List must have thought that my question was completely pointless because the dynamic analysis of conjunction was so obviously wrong. It was these reactions, in fact, which convinced me that this *is* obvious.

rationale for such an encoding? One might just as well hold that there is an English word which, as part of its lexical meaning, represents the principle that one should not interrupt a speaker who is telling a good joke, or that the use of a hearing aid may improve the quality of speech perception.

Consider the poor boxer who, due to an unfortunate accident which put an end to his career, processes incoming utterances from right to left. Must we say that it is impossible for him to understand the meaning of *and*, or that, if he does understand it, his linguistic behaviour is inconsistent with his linguistic knowledge? I fail to see that we are forced to conclude anything of the sort. But then the meaning of *and* is not dynamic.

On the standard truth-functional view, the lexical meaning of *and* is commutative. On this view, a child learning English does not have to worry about the order in which the members of a conjunction arise, insofar as his concern is with the interpretative effects of *and*. On the dynamic view, however, a child learning the meaning of *and* will have to worry about this: he must decide whether '$S_1$ and $S_2$' is to be interpreted as $[S_1] \circ [S_2]$ or $[S_2] \circ [S_1]$. This leads us to expect that in principle a child might adopt the working hypothesis that '$S_1$ and $S_2$' means $[S_2] \circ [S_1]$. Clearly, such considerations are too implausible to be taken seriously, but they *are* prompted by the dynamic approach, and its supporters are not in a position to reject them out of hand.

Could there be a language featuring a particle, *ugh* say, whose meaning is the mirror image of English *and*, i.e. a language in which '$S_1$ ugh $S_2$' is interpreted as $[S_2] \circ [S_1]$ ? It seems pretty evident that this would be a linguistic monster: there are no such languages, and linguistically speaking, there couldn't be any, either. However, if the dynamic analysis of conjunction is correct, then *and* and *ugh* are so close together in conceptual space that we must ask why it is that *ugh* is an impossible word, linguistically speaking. The question is absurd, it seems to me, but again the dynamic view on meaning implies that it deserves to be taken seriously.

Much to my surprise, the foregoing argument has been misunderstood more than once, so at the risk of belabouring the obvious, let me say a bit more about it. Suppose someone was to claim that the meaning of *and* is a truth function which yields 'true' if its arguments are both true, and 'false' otherwise. As this is not the only truth function, this claim would naturally prompt the question if there are any languages in which other truth functions are lexicalized. This is a good question, and it turns out that there are such languages. In fact, there are many such languages, English being one of them, as witness *or* and *neither ... nor* and, perhaps, *not*. Compare this with the claim that '$S_1$ and $S_2$' means $[S_1] \circ [S_2]$. This claim prompts the question if there are any languages with a connective that maps $S_1$ and $S_2$ (in this order) onto $[S_2] \circ [S_1]$. This does not strike me as a very good question. Frankly, I think it is silly question. The issue of *ugh* is on a par with that of the child ignoring the hypothesis that '$S_1$ and $S_2$' means $[S_2] \circ [S_1]$: they are not to be

taken seriously. Just as a pupil who asks a silly question will be the laughing stock of his class, a theory that gives rise to such non-issues is reduced, if not *ad absurdum*, then surely *ad ridiculum*.

I don't wish to deny that on the face of it the dynamic account of *and* has a certain appeal, but whatever initial appeal the theory may have derives from the fact that language is processed incrementally. As far as I can see, the only coherent construal of dynamic semantics is that it is an attempt at hardwiring this processing strategy into the lexicon.[2] But to grant that this is a coherent construal is not to condone the project. Indeed, if one's objective is to account for the observation that a sentence of the form '$S_1$ and $S_2$', like *any* complex expression, is processed from left to right, then the lexical entry of *and* surely is one of the most unlikely starting points. Not only that, but since the processing strategy is a perfectly general one, any such account is bound to miss an obvious generalization by a long shot.

In this section I have presented my principal argument against dynamic semantics. I believe to have shown that the dynamic analysis of *and* is wrong, and dynamic semantics stands or falls with this analysis. In the following I will adduce additional evidence against the dynamic approach, but the main point has been made.

## 4.2 Dynamic semantics vs. DRT

In the foregoing I tacitly presupposed Groenendijk and Stokhof's definition of dynamic semantics:

> A semantics is dynamic if and only if its notion of conjunction is dynamic, and hence non-commutative. (Groenendijk and Stokhof 1989)

This is stipulative, of course. 'Dynamic semantics' is a term of art, which we may define any way we want, but some definitions are more illuminating than others, and I feel that this one characterizes the dynamic approach rather well. Let me explain why.

An utterance changes the context in which it is made; utterances are acts, and acts have causal effects. It is this insight, evidently, that inspires the dynamic view on meaning, but inspiration is not the same as justification. The crucial tenet of dynamic semantics, after all, is that *sentences* are context change devices. This may or may not be correct, but it doesn't follow from

---

[2] Linear order has all sorts of interpretative effects, but dynamic semantics concentrates its attention on anaphoric processes, and the idea seems to be that the dynamics of *and* is only relevant to the interaction between linear order and anaphora. Strictly speaking, therefore, dynamic *and* is a *partial* lexical encoding of a processing strategy. (Not that this makes the dynamic account any more plausible. On the contrary.)

the observation that *utterances* change the context. However this may be, one premise of dynamic semantics is that sentences are context change devices, or, in Heim's terminology, that the meaning of a sentence is given by its context change potential, or CCP for short. A further one is that 'the CCPs of complex sentences can be given compositionally on the basis of the CCPs of their constituents.' (Heim 1983) To see how these two assumptions naturally, though not necessarily, give rise to the claim that conjunction is dynamic, compare the following examples with (2a, b):

(3)     a.    [A man]$_i$ walks in the park. He$_i$ whistles.

        b.    He$_i$ whistles. [A man]$_i$ walks in the park.

The contrast is clearly the same as in the earlier example: in (3a) anaphora is possible, and in (3b) it isn't. We should expect, therefore, that an explanation of the contrast in (3) will apply to (2), as well. In a dynamic semantics perspective, the difference between (3a) and (3b) is the following. In (3a) the first sentence sets up a context with an antecedent for the pronoun in the second one. In (3b), on the other hand, the pronoun occurs in a sentence that is supposed to set up a context for a sentence with the pronoun's antecedent, which is impossible. But if we want to extend this account to (2a, b), assuming that the CCP of a conjunction is determined compositionally by the CCPs of its parts, then it is hard to avoid the conclusion that the CCP of *and* is dynamic.[3]

So, although technically speaking it is possible to have a classical, 'static' conjunction in a theory like DPL, the claim that conjunction is dynamic follows naturally from the view, which is characteristic of the dynamic approach, that 'certain aspects of language use enter directly into the compositional core of a semantic system', as Chierchia (1995: xiii) puts it. Therefore, Groenendijk and Stokhof's definition of dynamic semantics is quite appropriate.

Another reason why I find this definition apt is that it draws attention to the fact that in a dynamic semantics ordinary connectives and quantifiers carry an inordinately heavy explanatory burden. The contrast between (2a) and (2b), for example, is explained in terms of the meaning of *and*. More generally, the dynamic research programme is to explain pragmatic

---

[3]   Thus the parallel between the examples in (2) and (3) is used to motivate a dynamic construal of *and*. However, as several people have pointed to me, it is much more plausible to interpret this parallel as demonstrating, on the contrary, that the meaning of *and* cannot be dynamic. If the meaning of *and* is involved in the explanation of contrast in (2), then how are we to account for the contrast in (3), where there is no linguistic symbol that can be taken to denote '∘' ? Inevitably, if he wants to save the insight that (2) and (3) are alike in all relevant respects, the dynamic semanticist is forced to maintain that sentence boundaries *mean* '∘' in the same sense in which *and* means '∘'. Indeed, it has been claimed that the full stop denotes '∘'. I am not making this up: see Muskens (1990), for example.

phenomena such as anaphora and presupposition in lexical terms; especially, in terms of the lexical meanings of *and*, *or*, *not*, etc.

It is customary to bracket together DRT with such dynamic theories as DPL. As I have pointed out already this is unfortunate if not downright misleading, because DRT is not a dynamic semantics. DRT is a theory of language understanding; as Kamp and Reyle say, DRT aims to provide:

> [...] an (idealized) analysis of the process whereby the recipient of an utterance comes to grasp the thoughts that the utterance contains. We have been led towards our description of the relation between meaning and form by a firm conviction that it is only in this way that a comprehensive range of semantic facts can be explained. (Kamp and Reyle 1993: 8)

This theory of language understanding is a representational one in the sense that it describes how a hearer constructs mental representations, called 'DRSs', in response to linguistic stimuli. The concept of context change that lies at the heart of dynamic semantics is foreign to DRT.

In a dynamic semantics sentences are viewed as context change devices. It is by no means obvious how this is to be understood, as sentences do not change contexts any more than shoes, ships, or sealing wax do. Speakers can *use* sentences for making utterances, and the objective of DRT is to explain how hearers respond to such utterances, focusing on the mental representations that they construct in the process. Might we say, then, that in DRT contexts are represented by DRSs, and sentences are viewed as context change devices, after all, because they serve to modify DRSs? We might, of course, but this construal would be perverse. An agent can use a sentence in order to provoke changes in his fellow agents' mental representations, but in principle he can use anything for that purpose (and often non-linguistic means will prove to be more effective). So if a sentence counts as a context change device just because it may be used to affect mental representations, *anything* is a context change device.

DRT is a non-compositional theory of meaning. This is not to say that it cannot be reformulated in strictly compositional terms, for it has been known since Zeevat (1989a) that this can be done. It is just to say that compositionality of meaning doesn't play a role in DRT; it is optional. In my opinion this is as it should be, because I don't know of any principled reason why a semantic theory should be compositional, but dynamic semanticists are wont to claim that compositionality is one of the prime virtues a semantic theory can aspire to, and to chastise DRT for being non-compositional. However, I don't want to discuss the compositionality issue as such (as far as I can tell there isn't much of an issue anyway); I just want to point out that the emphasis laid on compositionality by dynamic semanticists shows that their views on semantic theory are quite different from the position taken by

DRT. In particular, it shows that DRT's breakaway from the Montagovian tradition is much more radical than that of dynamic semantics.

DRT is a representational theory of meaning. It is a theory about the mental representations that hearers construct in response to utterances. In contrast, dynamic semantics purports to be a non-representational theory. It is not all too clear what this is supposed to mean. I should say that a variable assignment contains linguistic, hence representational, elements, and that a theory which regards variable assignments as semantic objects is *ipso facto* a representational theory. But that is as it may be. I am more interested in what fuels the controversy over representations in semantics. One is often given the impression that this discussion is about how many 'levels' are needed in a linguistic theory. But really the issue is more fundamental than, say, the debate over deep structures in syntax. What is at stake is the status of semantic theory. If it is true, as Kamp and Reyle claim, that semantic facts can only be explained by way of a theory of language understanding, and if we discount the unlikely position that a non-representational theory of understanding is a feasible project, then arguments against semantic representations are in fact arguments against semantics as a theory of understanding. I suspect, therefore, that dynamic semanticists' qualms about representations betray their dislike of the philosophical underpinning of DRT. What foments this dislike, presumably, is that dynamic semantics is a realist theory of meaning. This realism is veiled by talk of information and information exchange, common ground, and context change, but it is unmistakably there. According to the semantic realist, meaning is a relation between language and the world. The dynamic semanticist accepts this position, too; he only extends the realist's standard repertory of semantic values with such beasts as context change potentials, relations between assignments, and so on.

To sum up: DRT differs from dynamic theories of meaning not only because compositionality is not on its list of desiderata, but also, and more importantly, because it is a theory of language understanding, not of context change, and hence a representational theory. That DRT lies squarely outside the dynamic paradigm also appears from the fact that its supporters do not share the preoccupation with connectives that is symptomatic of dynamic semantics. Within the DRT camp, there have never been serious controversies over the meaning of *not* or *or*, for example, and it is fairly obvious what a DRT account of conjunction should look like.

It is admittedly unfortunate that what I take to be the correct DRT account of conjunction does not coincide with the analysis espoused by Kamp and Reyle (1993). In Kamp and Reyle's version of DRT, sentences of the form '$S_1$ and $S_2$' are handled by first extending the current DRS with the information in $S_1$, which results in a new DRS, in the context of which $S_2$ is interpreted. This is what one should expect, of course. What is surprising is that, in their

implementation of this analysis, Kamp and Reyle adopt a purely lexicalist approach: they set up a lexical rule associated with *and*, which, by means of an indexing scheme introduced expressly for this purpose, ensures that $S_1$ is processed before $S_2$. Clearly, this analysis is very much in the spirit of dynamic semantics, and all objections against the dynamic treatment of conjunction apply to Kamp and Reyle's proposal, too. However, within the DRT framework their treatment is anything but inevitable.

I have already hinted at what I take to be the right analysis of conjunction. It is the orthodox analysis, according to which the lexical content of *and* merely says that the sentences flanking it are both true. In terms of Kamp and Reyle's version of DRT, this is just to say that the lexical rule associated with *and* is an instruction to interpret both conjuncts within the context of the current DRS. Hence, as far as the lexical meaning of *and* is concerned, conjunction is commutative. If the order of the members of a conjoined sentence matters, as it often does, it is not because the lexical meaning of *and* is non-commutative, but because the process of interpretation is sensitive to the order in which it receives its input, as in general it has a preference for incremental processing.

This is just a sketch, to be sure, but it can be made quite precise by borrowing concepts from computer science, more especially from parsing theory. The construction rules of a system *à la* Kamp and Reyle may be viewed as a grammar, which is applied by a semantic parser, as one might say. Formulated in these terms, the notion of incremental interpretation may be defined by saying that, by default, the semantic parser processes its input depth-first and left-to-right.

## 4.3 Forward reference

Whereas I say that incremental interpretation is just a processing strategy, the dynamic semanticist says that it is encoded in the meaning of *and*. Our positions yield different predictions about the interpretative effects of linear order. If we adopt the DRT analysis that I advocate, it should be possible for an anaphoric expression in the first half of a conjunction to have its antecedent in the second conjunct; for the expectation that this will not happen is based upon a processing strategy, and is therefore defeasible. If, on the other hand, the dynamic semanticist is right, then this should be impossible, since on his analysis it follows from the lexical meaning of *and* that a pronoun in the first conjunct cannot have its antecedent in the second conjunct. It is the former prediction rather than the latter which is borne out by observations like the following, which I borrow from Bolinger (1977):

(4)      a.     Yes, the doctor warned him, but Tom kept on.

         b.     He looks at me and John goes out of his mind.

         c.     He lied to me, and John was my friend!

In each of these examples an anaphoric pronoun in the first conjunct refers forward to an antecedent in the second. Intuitively, such sentences are most likely to be uttered in a context in which the intended antecedents of these pronouns have already been introduced, but there are at least two ways of showing that such anaphors cannot always be construed as surreptitiously referring backwards. First, consider the following discourse:[4]

(5)      Barney and Fred were arguing when I came in. Well, I'm standing by the door and then *he* looks at me and {he/Fred} goes out of his mind.

Apparently, the italicized occurrence of *he* is at least dependent on the second occurrence of *Fred*; for if we replace the latter with a pronoun, it becomes unclear whom the first *he* is intended to refer to. This should suffice to dispel the simple-minded view that forward reference is just an underhanded way of referring backwards. The same point is brought home more forcefully by the observation that forward referring pronouns may have indefinite antecedents, as the following examples show:

(6)      a.     I don't know what $he_i$ has on $them_j$, but it seems that [one of the pupils]$_i$ is blackmailing [some of the teachers]$_j$.

         b.     I don't know what $they_i$ have on $them_j$, but it seems that [most of the pupils]$_i$ are blackmailing [at least one or two of the teachers]$_j$.

In (6a), *he* and *them* refer forward to *one of the pupils* and *some of the teachers*, respectively. These indefinites are most likely to be construed specifically, but as (6b) demonstrates, this is not a prerequisite for forward reference: in this example it is hard if not impossible to construe *at least one or two of the teachers* specifically, but nonetheless it is the intended antecedent of the preceding pronoun *them*.

These observations are at odds with the dynamic account of anaphoric reference in conjunctions, and in particular cases like (6a, b) are counter-examples against the standard dynamic approach. The question therefore is if the dynamic theory can be repaired so as to capture these data. I can see a

---

[4]   Here and in the following I am indebted to van Deemter's (1990) discussion of forward reference.

number of possible repairs, none of which are particularly successful, though.[5]

(*i*) Someone steeped in the principles of dynamic semantics will almost instinctively start tinkering with the definitions of the connectives, in this case *and*. For instance, he might consider inverting the standard definition of the semicolon, as follows:

(7)      $a[\varphi \vdots \psi]b$ iff $\exists c: a[\psi]c$ and $c[\varphi]b$, i.e. $\varphi \vdots \psi \Leftrightarrow \psi ; \varphi$

Truth-conditionally speaking ' $\vdots$ ' is equivalent with ' ; ': it merely reverses the order in which the two members of a conjunction are processed. Hence, on the assumption that in (6a, b) *and* symbolizes ' $\vdots$ ', these examples can now be accounted for. However, since in most cases pronouns follow their antecedents, it will have to be claimed that *and* is actually ambiguous between ' $\vdots$ ' and ' ; '.

There are several problems with this solution. First, having argued at some length that for various reasons it is hard to accept that conjunction is dynamic, I don't see that these difficulties are alleviated by the addendum that it is dynamically ambiguous. Secondly, it will clearly not do to postulate that conjunction is ambiguous, because normally speaking forward reference across *and* is not possible. So apart from the fact that the lexical entry of *and* would have to explicitly allow for a conjunction to be processed either from left to right or from right to left, it would also have to specify that the former option is generally preferred to the latter. The same objection now applies with a vengeance: there is no reason whatsoever to suppose that this processing information is lexically encoded. And not only is this supposition unfounded, it results in a cumbersome account of the facts. Thirdly, if *and* really were ambiguous between ' $\vdots$ ' and ' ; ', we should expect there to be languages which have distinct morphemes for these two meanings. I don't think it is very likely that such a language will ever be found. Finally, this proposal fails to account for conjunctions with crossing anaphoric dependencies, such as the following:

(8)      a.      Fred worked on it$_i$ [for over a year]$_j$, and then$_j$ he emerged from his workshop with [one of the most beautiful ukuleles the world has ever seen]$_i$.

---

[5] Forward reference in conjunctions is to be distinguished from forward reference in certain other constructions, like conditionals:

> When it$_i$ is hungry, a cat$_i$ usually meows.
> If he$_i$ lies to a student$_j$, a teacher$_i$ loses his$_j$ trust.

Chierchia (1995) discusses these examples in a dynamic framework, but the treatment he proposes does not apply to conjunctions, and it is unlikely that it can be extended so as to deal with the phenomena under discussion. Cf. below, under (*ii*).

     b.    It still is a mystery how she managed to smuggle them$_i$ into [her apartment]$_j$ without being noticed, and yet it is now established beyond any reasonable doubt that Betty has been keeping [twelve German shepherds]$_i$ there$_j$ for more than a decade.

In (8a), the pronoun in the first conjunct refers forward to the indefinite NP in the second, while at the same time *then* in the second conjunct refers backwards to the temporal adverbial in the first; and the same holds, mutatis mutandis, for the pronouns *them* and *there* in (8b).

    It might be suggested that such examples can be handled by yet another version of dynamic conjunction, along the following lines:

(9)      $a[\varphi \, ; \, \psi]b$ iff $\exists c$: either $a[\varphi]c$ and $c[\psi]b$, or $a[\psi]c$ and $c[\varphi]b$

This will not do, however, because it doesn't allow for crossing anaphoric links; it only allows for anaphoric links to collectively run in one direction or the other.

    (*ii*) A rather different approach to forward reference is to claim that it is a matter of scope: whenever a pronoun is used to refer forward, its antecedent is construed as taking scope over the clause in which the pronoun occurs. Thus in (6b), for example, *most of the pupils* and *at least one or two of the teachers* are construed as having widest and second-widest scope, respectively. Technically speaking, this proposal is unproblematic. There are several methods for handling scope, and in principle any of these might be pressed into service for dealing with forward reference. But apart from its technical feasibility this approach has little to recommend it.

    First, it is at odds with the widely accepted premise that there are certain grammatical constraints on scope taking: if we accept that (6b), for example, must be analysed in terms of scope, then apparently we are forced to conclude that, with respect to scope taking, practically anything goes, which is surely incorrect.

    Secondly, the scope analysis goes against the spirit of dynamic semantics. To say that natural language conjunction is dynamic is to say that it is processed from left to right, i.e. that the second conjunct is processed in the context of the first. But, evidently, if an NP in the second conjunct is allowed to outscope the first, then the first conjunct is interpreted in the context of (part of) the second. Even if this is unproblematic from a technical point of view, it is hard to fathom what the explanatory merit of such an account might be.

    Thirdly, if a dynamic semanticist decides to advocate a scope analysis of forward reference, he will be pulling the rug from under his own theory. For if we can account for forward reference along these lines, then surely we can treat *all* sorts of anaphora in conjoined sentences in terms of scope, without

having to assume that conjunction is dynamic. In fact, a static theory of scope can handle practically anything a dynamic semantics can account for.[6] But then there is no point in maintaining that the semantics of natural language is a dynamic one.

(*iii*) All remaining options are, as far as I can see, quite hopeless. It has been suggested to me, for example, that forward reference is a marginal phenomenon that shouldn't be given so much weight, or that it should be relegated to a pragmatic theory (!). Obviously, these are counsels of despair, and I will not discuss them any further.

The long and short of it is that forward reference stands as clear evidence against the theory of dynamic semantics. To a theory like DRT, forward reference does not present any fundamental difficulties. This is not to suggest that, as it stands, DRT says all there is to say about forward reference. It is just to say that if one adopts DRT at the least the outlines of an explanation are given, and it seems to be a plausible explanation, too. Whereas, if one adopts the dynamic perspective, the best strategy for dealing with forward reference is to shrug one's shoulders and turn away.

To conclude this section, let me repeat that I brought up this subject merely to amplify the conclusion which was established already in § 4.1. It is not my position that dynamic semantics is a hopeless enterprise because it cannot deal with forward reference. My position is, rather, that dynamic semantics is hopeless because it mistakes processing factors for aspects of meaning. Its inability to deal with forward reference merely confirms this assessment; it is not an indispensible piece of evidence.

## 4.4 Other connectives, same story

Thus far I have concentrated my attention on *and*, whose analysis epitomizes the dynamic approach. In this section, I want to show that the points I have made in the foregoing carry over to the other connectives, in particular *not* and *or*. In all varieties of dynamic semantics, certain laws of classical logic cease to hold. In particular, the law of double negation tends to break down in a dynamic setting, and while in classical logic (10) is valid,

(10)     $\neg\varphi \vee \psi \Leftrightarrow \varphi \rightarrow \psi$

---

[6] This is shown by Dekker (1993). To remain on the safe side, I say 'practically anything' because Dekker maintains that there are certain phenomena that must be treated by construing negation as a dynamic operator, and such phenomena cannot always be accounted for in terms of scope. Since it is doubtful, however, that there is any need for a dynamic construal of negation (see below), this subtlety may be ignored.

this equivalence does not obtain in the more popular versions of dynamic semantics, such as DPL for example. With these observations in mind, let us turn to the following examples:

(11)    a.    It's not true that this house doesn't have a bathroom. It's on the second floor.

        b.    Either there is no bathroom in this house or it is in a funny place.

In DPL, the pronoun *it* in the second sentence of (11a) cannot have *a bathroom* as its antecedent, because negated sentences are tests: an antecedent introduced within the scope of a negation cannot be picked up from outside. This prediction surely holds for simple cases, but data like (11a) suggest that double negation should make an antecedent accessible again. If the law of double negation were to hold in a dynamic system like DPL, we would have an explanation for the well-formedness of (11a). Similar considerations apply to (11b). In DPL, as in DRT, a pronoun on one side of *or* cannot have its antecedent on the other side, because disjunctions are internally static. Again, in general this is correct, but (11b) would appear to be an exception to the rule; for if the law of double negation and (10) were to hold in a DPL-type system, we would have an explanation for cases like (11b), too.

   Since I don't want to get bogged down in technicalities, and the points I want to raise can be formulated without going into the logical details, let us simply assume that we have a dynamic semantics, call it $DPL^+$, which meets these requirements. That is, in $DPL^+$ (10) holds and double negation cancels out. Dynamic systems with these properties have been developed by Dekker (1993) and Krahmer and Muskens (1995), so the assumption that $DPL^+$ might exist is justified. In $DPL^+$ (11a, b) are represented as (12a) and (12b), respectively:

(12)    a.    $\neg\neg\exists x Px \; ; Qx \Leftrightarrow \exists x Px \; ; Qx \Leftrightarrow \exists x(Px \; ; Qx)$

        b.    $\neg\exists x Px \lor Qx \Leftrightarrow \exists x Px \to Qx \Leftrightarrow \forall x(Px \to Qx)$

Hence, $DPL^+$ predicts that these sentences are equivalent with (13a) and (13b), respectively.

(13)    a.    This house has a bathroom, and it's on the second floor.

        b.    If there is a bathroom in this house it is in a funny place

This seems to be correct.

   Still, this analysis is arguably on the wrong track. To start with, the $DPL^+$ treatment of double negation covers only a fraction of what is in fact a much more general phenomenon. To illustrate, consider:

(14)    It's ludicrous to pretend that this house doesn't have a bathroom. You showed it to me, remember?

Here anaphora is possible, and this example is obviously of the same making as (11a). However, since DPL$^+$ is designed to account for the double-negation cases only, it doesn't have an explanation for examples like (14). As I argued in § 2.4.3, it is clear at least in outline how such cases should be analysed. If someone utters the first sentence of (14), then the hearer may *infer* that according to the speaker the house has a bathroom. So the anaphor in the second sentence is accounted for on the assumption that, under certain circumstances, such inferential processes may give rise to new reference markers that can be picked up by subsequent anaphoric pronouns. In other words, these anaphors are interpreted by way of bridging. This type of anaphora may seem to be rare, but examples are easier to find than one might think:

(15)　　a.　　The problem is not that I don't know what to say. It's how to say *it*.

　　　　b.　　You're not a true Russellian unless you accept *his* views on marriage and education, too.

　　　　c.　　In Amsterdam *they* don't believe in DRT anymore.

　　　　d.　　If you hear strange sounds coming from the basement, *it*'s probably Fred working on his oboe exercises.

　　　　e.　　He wants to become an oboist because he thinks *it*'s such a lovely instrument.

　　　　f.　　I don't wish to deny that I own a donkey. But I don't beat *it*.

　　　　g.　　Every time the dog relieves itself on the carpet, I have to clean *it* up.

Although I readily admit that it is by no means clear what the details of a bridging analysis will look like, it is clear that an approach along these lines is called for if we want to handle examples like these. But if a bridging account must be adopted for the examples in (14) and (15) anyway, then why should one resort to a different analysis for the double-negation cases? They can be treated as instances of bridging anaphora, too, so there is no need to deploy any special-purpose machinery. On the contrary, any theory that does so disqualifies itself on methodological grounds.

　This problem is not restricted to examples like (14). Here is another pair of problematic cases:[7]

(16)　　a.　　John remembered to bring an umbrella. It's in the hallway.

　　　　b.　　John didn't forget to bring an umbrella. It's in the hallway.

---

[7]　Krahmer and Muskens (1995) mention this pair in a footnote, attributing the crucial observation to an anonymous referee for the *Journal of Semantics*.

DPL$^+$ doesn't say anything about (16a): there is no connective in sight that could be proclaimed 'dynamic', and so it fails to explain why anaphora is possible in this case. (16b) makes things even worse. Even if we are prepared to accept a decompositional analysis on which *forget = not remember*, as Krahmer and Muskens (1995) propose, DPL$^+$ just predicts that (16b) is the same as (16a), but this doesn't solve the problem. So in both cases a bridging account must be resorted to.

Yet another problem is caused by the following contrast:

(17)     It is not true that there wasn't a single guest at Betty's party.

   a.    ?But he was out on the terrace.

   b.    But they were all out on the terrace.

A bridging account easily explains this contrast: on the basis of world knowledge about parties it may be inferred from the first part of (17) that there was more than one guest at Betty's party, and therefore a plural pronoun is called for. In DPL$^+$, however, the logic eliminates the double negation, thus predicting that the first sentence of (17) is equivalent to — well, equivalent to what? To 'There was a single guest at Betty's party', perhaps? But then (17b) should be awkward, not (17a). At any rate, not only is it otiose to 'make negation dynamic'; it actually creates problems instead of solving them.[8]

What I have said about double negation carries over to the bathroom disjunctions, but these cause additional problems of their own. Consider the following contrast:

(18)     a.    Either there is no bathroom in this house or it is in a funny place. (= (11b))

   b.    Either it is in a funny place or there is no bathroom in this house.

(18a), in which the pronoun follows its antecedent, is a fairly normal sentence. By contrast, (18b), in which the order of the disjuncts is reversed, is decidedly marked. Let us grant for argument's sake that the DPL$^+$ account of negation is correct, and that, given the equivalences in (12b), the theory's interpretation of '∨' accounts for (18a). But then (18b) is ruled out. This seems too drastic however. Although (18b) is marked, it is perfectly intelligible, and just the kind of the thing that someone might say under certain circumstances. Note, furthermore, that if we replace the pronoun with a full definite NP, the order of the disjuncts does not make much of a difference anymore:

---

[8]   Krahmer and Muskens remark, apropos of an analogous example, that the difference is due to a 'uniqueness effect' (1995: 359). They don't explain what they mean by this.
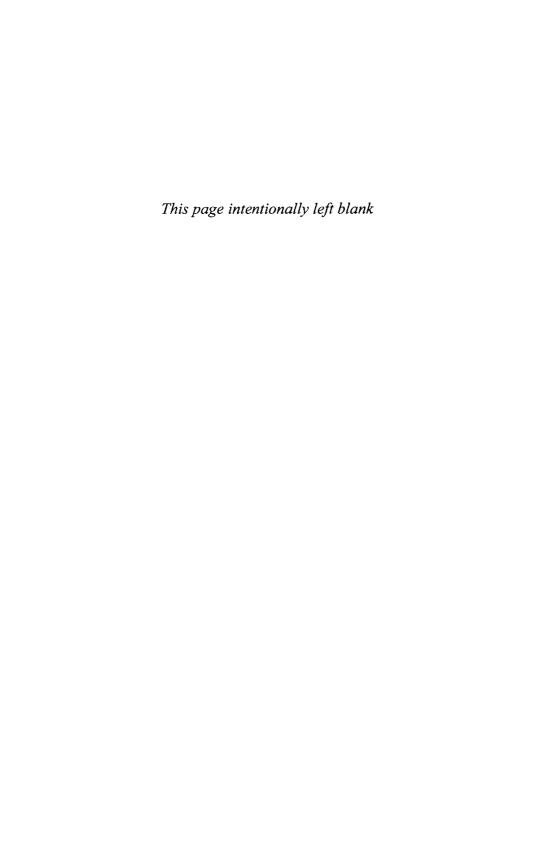
(19)  a.    Either there is no bathroom in this house or it the bathroom in a funny place.

b.    Either the bathroom is in a funny place or there is no bathroom in this house.

Of course, this problem is essentially the same as the one discussed in the previous section. We have to account for the fact that, in principle, forward reference is possible, although in general anaphors prefer not to precede their antecedents, and the lexicalist stance taken by dynamic semanticists is not adequate to this purpose.

## 4.5 Conclusion

It is evident that linear order may affect the interpretation of an utterance. The most obvious explanation of this fact is also the simplest and the most adequate one: it is just that hearers have a preference for incremental interpretation. One of the central tenets of dynamic semantics is that certain effects of linear order are encoded in the lexical meanings of the connectives, in particular of *and*. This assumption is redundant, it is implausible, and it makes false predictions. Dynamic semantics is not just wrong: it is a nonstarter.

In the preceding chapter we have seen that the satisfaction theory of presupposition is fatally flawed, and in this chapter it has been shown that the same holds for the dynamic framework which that theory presupposes. These results lend additional support to the binding theory and the DRT framework in which it is couched. Not only is the cognitivist approach to presupposition and anaphora successful in its own right: the competition is not exactly fierce, either.

*This page intentionally left blank*

CHAPTER 5

# Attitude reports

Although the literature on attitude reports is probably even larger than the presupposition literature, surprisingly little ink has been spilled on the interaction between presupposition inducers and attitude verbs. I don't have an explanation for this discrepancy, but it is certainly not as if the problems posed by presuppositions triggered in attitude contexts weren't challenging enough. For, as will be show in § 5.1, such presuppositions give rise to a number of perplexing puzzles. In § 5.2, I turn to Heim's (1992) theory, which is an attempt to solve some of these puzzles, and criticize her solution on several counts. In the remainder of this chapter I present my own proposal. To begin with, I define an extension of DRT in which attitude reports can be represented (§§ 5.3 and 5.4). I then argue that the verb *believe* triggers a certain presupposition and that this presupposition determines the projection behaviour of the verb (§ 5.5). This analysis yields predictions which are incomplete in the sense that they do not fully account for speakers' intuitions about presuppositions and anaphors in belief contexts. In § 5.6, I suggest that this is as it should be because we may assume on independent grounds that there is a class of plausibility inferences which complement the predictions of the binding theory. In the final section, the analysis is extended to the verb *want* (§ 5.7).

## 5.1  Presuppositional expressions in attitude contexts

There are at least two reasons why presuppositions that are triggered within the scope of an attitude verb are more problematic than others. One is that attitude contexts are a notoriously difficult area in their own right. The other is that in attitude contexts the pre-theoretical notion of presupposition, which otherwise is quite secure, starts tottering. Consider a simple example:

(1)      Louise believes that her niece lives in Leeds.

In (1) the definite description *her niece* occurs within the scope of the attitude verb *believe*. We would normally be entitled to infer from an utterance of this sentence that (the speaker believes that) Louise has a niece, and it may be

argued that this inference is of a presuppositional nature. Generalizing this observation, we arrive at

(2)     *The e-principle*

    If $V$ is a verb of propositional attitude and $S\{\chi\}$ is a simple sentence in which the presupposition is triggered that $\chi$, then (a speaker who utters) '$a$ $V$s that $S\{\chi\}$' presupposes that $\chi$.

I call this the 'e-principle' because it says that a presupposition that arises in the attitude context associated with $V$ is construed externally, i.e. as not being a proper part of the attitude report. There may be some temptation to view this as a special kind of *de re* construal, but as I shall argue below, what I call the external construal of a presupposition is sufficiently different from the notion of *de re* interpretation to deserve a name of its own. Construed as a rough empirical generalization, not as a theoretical statement, the e-principle appears to inform Gazdar's (1979) and van der Sandt's (1988) views on the interaction between presuppositions and attitude contexts, and the principle looks plausible enough, at least as long as we are prepared to ignore the fact that (1) would usually license a *further* inference, which also seems to originate with the definite noun phrase *her niece*, namely that *Louise believes* that she has a niece. Hence the following generalization might be entertained as well:

(3)     *The i-principle*

    If $V$ is a verb of propositional attitude and $S\{\chi\}$ is a simple sentence in which the presupposition is triggered that $\chi$, then (a speaker who utters) '$a$ $V$s that $S\{\chi\}$' presupposes that $a$ believes that $\chi$.

Terminological details aside, the i-principle ('i' for 'internal') is defended by Karttunen (1974) as well as Heim (1992). Again, the concept of internal construal is similar to, but not quite the same as, the concept of *de dicto* construal. The difference (which will be discussed in greater detail in § 5.2) is already apparent from the wording of (3), which says that '$a$ $V$s that $S\{\chi\}$' presupposes, not that $a$ $V$s that $\chi$, but rather that *a believes* that $\chi$. To see why this is so, consider:

(4)     Louise $\left\{ \begin{array}{c} \text{doubts} \\ \text{suspects} \\ \text{hopes} \end{array} \right\}$ that her niece lives in Leeds.

As noted already by Karttunen (1974), an utterance of (4) would not normally be taken to imply that Louise doubts/suspects/hopes that she has a niece (which is what we would expect if *her niece* were read *de dicto*), but rather that she believes that she has one. Among the attitudes, belief has a special status: our doubts, suspicions, and hopes are grounded in what we

believe, but not, or not necessarily, the other way round. It is evident that the regularity expressed by the i-principle has something to do with this asymmetry, although it remains to be seen what exactly is the nature of the connection.

The problem with (1) and many other attitude ascriptions, is that they may reasonably be taken to corroborate both the i-principle and the e-principle. We would normally infer from an utterance of (1) that Louise has a niece *and* that she believes that she has one — or, to put it otherwise, presuppositions that are triggered in the scope of an attitude verb may give rise to a two-sided reading, which is internal as well as external.[1] Such two-sided readings do not always come about, of course. In a situation in which Louise doesn't know that she has a niece, but is nonetheless acquainted with her niece in some other way, a merely external reading of (1) would be called for; whereas in a context in which Louise mistakenly believes that she has a niece, a strictly internal reading would be achieved. But often a two-sided reading will be the most natural one.

It may be argued that both the internal and the external halves of these two-sided interpretations are of a presuppositional nature. First, it is clear that both the inference that Louise has a niece and the inference that she believes that she has a niece can be traced back to the definite noun phrase *her niece*. Secondly, both inferences appear to display the projection behaviour that is the hallmark of presuppositions.

(5)     Maybe Louise doesn't believe that her niece lives in Leeds.

Here (1) is embedded within the scope of a modal and a negation operator, and as one would expect on the basis of the e-principle, it follows from (5) that Louise has a niece, while the i-principle correctly predicts that an utterance of this sentence would normally imply that Louise believes that she has a niece.

In the light of these observations, it may seem attractive to take the position that both the i-principle and the e-principle are correct. Things are not as simple as that, however. Suppose that we extend our DRS language in such a way that representations like the following are well formed:

(6)     [: Louise believes: [x: x is Louise's niece, x lives in Leeds]]

This is intended as a nearly completed representation of (1), in which only one presupposition remains to be dealt with, i.e. that Louise has a niece. The precise semantics of conditions of the form '$u$ believes: $\psi$' will be left open

---

[1]  Loar (1972) may have been the first one to observe two-sided readings, but his remarks are restricted to definite NPs. That other presupposition inducers can have such readings, too, was recognized, either implicitly or explicitly, by Fauconnier (1985), van der Sandt (1988), Heim (1992), Kay (1992), and Zeevat (1992).

for the time being, since this is precisely one of the issues at stake; for the moment it suffices if some intuitive sense can be made of representations like (6).

Although it is the semantic interpretation of belief conditions that will determine whether or not, in a DRS φ that contains a condition of the form '*u* believes: ψ', φ is accessible to ψ, it is instructive to pretend that we may simply decide the matter by fiat, and see what happens if we choose one way or the other. So let us assume that the main DRS in (6) is accessible to the embedded one. Then the binding theory predicts that (6) will be completed as follows:

(7)      [x: x is Louise's niece, Louise believes: [: x lives in Leeds]]

Here Louise's niece has been accommodated in the main DRS. Hence, if the clausal complement of *believe* has access to the context in which it is embedded, we confirm the e-principle but not the i-principle.

On the other hand, suppose that our semantics entails that the main DRS in (6) is not accessible to the embedded DRS. We then predict that Louise's niece must be accommodated in Louise's belief DRS. Hence, if the embedded DRS in (6) does not have access to the main DRS, the binding theory predicts an internal instead of an external reading, and the i-principle is confirmed, not the e-principle. So it is far from obvious how we may obtain two-sided readings if we start out from the premise that both the internal and the external aspect are of a presuppositional nature, and must therefore both be accounted for by a theory of presupposition.

At this point another strategy suggests itself. Suppose that we have a semantics that allows us to make sense of a DRS like (7), which is to say that the main DRS is accessible to the embedded one. As it stands, the binding theory will then predict that (1) must in fact have the interpretation encoded by (7). But now we add to our theory a further stipulation to the effect that every presupposition $\chi$ must be copied into every DRS between the DRS in which $\chi$ was triggered and the DRS to which it is projected. Implementational details aside, this is Zeevat's (1992) proposal. Zeevat's copying rule requires that (7) be transformed into:

(8)      [x: x is Louise's niece,
         Louise believes: [x: x is Louise's niece, x lives in Leeds]]

Thus it is predicted that (1) will receive the two-sided reading that we are after. Moreover, the resulting theory vindicates the i-principle as well as the e-principle, although it must be noted that the latter comes out as being logically prior to the former, for we first obtain the interpretation that the e-principle would predict, and then turn it into an interpretation that in addition satisfies the i-principle. However, the postulate that gives us this is quite clearly ad hoc, and lacks a sound conceptual foundation.

Presuppositions are supposed to be given pieces of information. What could ever be the rationale for requiring that the same piece of information be given more than once? Suppose that $\chi$ is a presupposition that is triggered in DRS $\varphi_1$ and bound in a higher DRS $\varphi_n$; $\varphi_1, \ldots, \varphi_n$ are the DRSs that are accessible to $\varphi_1$ and have access to $\varphi_n$. Once $\chi$ has been bound, it is established once and for all that $\chi$ was given, so why should a hearer take the trouble of placing copies of $\chi$ into all of $\varphi_1, \ldots, \varphi_{n-1}$? The same question may be asked if $\chi$ is accommodated, rather than bound, in $\varphi_n$. One should expect that accommodating $\chi$ once should suffice for upholding the pretence that $\chi$ was given, so why accommodate it, additionally, in each of $\varphi_1, \ldots, \varphi_{n-1}$? To the best of my knowledge, there isn't a satisfactory answer to these questions. Moreover, it seems that Zeevat's copying rule will make false predictions for attitude contexts created by verbs other than *believe* or its synonyms.[2] In the case of (4), for example, it would predict that Louise doubts/suspects/hopes that she has a niece, and this prediction is incorrect.

Although what I have said does not quite prove the point, it begins to look as if the i-principle and the e-principle cannot be true together. If the main DRS in (6) is accessible to the embedded one, the e-principle is confirmed but not the i-principle; if it is not accessible, it is the other way round; and Zeevat's attempt to have it both ways is ad hoc and yields incorrect predictions. It would seem, therefore, that in a representation like (8) — which is or at least resembles what we would like to have — at most one of the two sub-structures representing Louise's niece can be accounted for in presuppositional terms. The question then becomes whether a presupposition theory should explain the internal or the external half of a two-sided reading.[3] That is to say, should an account of the interaction between presuppositions and attitudes take its lead from the i-principle or should it rather take the e-principle as its point of departure?

In this chapter I consider two approaches to the problems that I have outlined in the foregoing. The first approach is advocated by Heim (1992), and may be characterized, with some qualifications, as an attempt to vindicate the i-principle: according to Heim, the inference from '*a* believes that $S\{\chi\}$' to '*a* believes that $\chi$' is of a presuppositional nature, while the inference to $\chi$ is not. I will argue in the next section that Heim's theory is on the wrong track, and then proceed to lay out an alternative approach which is similar to Heim's in some respects (in particular, it incorporates essentially the same semantics for *believe* and *want*), but which takes the e-principle as its starting point. I shall claim therefore that, as a rule, '*a* believes that $S\{\chi\}$'

---

[2] Zeevat (1992) confines his attention to belief contexts, and doesn't consider any other attitude verbs.

[3] I take it that the fourth possibility (that neither the internal nor the external component is explainable in presuppositional terms) is a non-starter.

presupposes χ, and that the inference to '*a* believes that χ' is to be accounted for in another way. One of my arguments in favour of this position is that such an account is in fact easy to provide, whereas it is much harder, if not impossible, to come by an inferential schema that runs in the opposite direction, which is what Heim requires.

Prima facie, it would seem that the e-principle is corroborated by (9a) and falsified by (9b), whereas with the i-principle it is the precisely the other way round:

(9) a. If Louise has a niece, then she believes that her niece lives in Leeds.

  b. If Louise believes that she has a niece, then she also believes that her niece lives in Leeds.

The i-principle predicts that in the consequent of both (9a) and (9b) the presupposition will be triggered that Louise believes that she has a niece. This presupposition will be bound in (9b), which is intuitively correct. However, in (9a) this presupposition will not be bound, so we predict that this sentence presupposes that Louise believes that she has a niece, which is false. In short, the i-principle gives the right prediction for (9b) but fails with (9a). Exactly the converse holds for the e-principle: it delivers the desired result for (9a), but also predicts that (9b) presupposes that Louise has a niece, which is just as false. One of my aims in this chapter will be to explain away this 'paradox'.

Another puzzle that I will try to solve is the following. Speaking in terms of the binding theory, (9a, b) suggest that a presupposition triggered in a belief context may be bound within as well as outside of that belief context. However, the data are more complex than this pair of examples suggests, as witness:

(10) a. Barney $\left\{\begin{array}{c}\text{believes}\\\text{hopes}\end{array}\right\}$ that Fred will agree, and he hopes that Wilma will agree too.

  b. ?Barney hopes that Fred will agree, and he believes that Wilma will agree, too.

(10a) shows that a presupposition triggered within a context created by the verb *hope* may be bound within a preceding *hope* or *believe* context, or within a *believe* context, and (10b) corroborates the asymmetry between *believe* and non-*believe* contexts that we observed earlier on, for apparently a presupposition triggered in a *believe* context cannot without further ado pick up an antecedent from a *hope* context.

The examples discussed in the foregoing are obviously related to the phenomenon of 'intentional identity', which has been an unfailing source of controversy ever since Geach published his famous Hob-Nob sentence:

(11)     Hob thinks a witch has blighted Bob's mare, and Nob wonders
         whether she (the same witch) killed Cob's sow. (Geach 1967: 628)

The problem of intentional identity as presented by Geach is the following.
It seems possible to read (11) in such a way that, according to the speaker,
Hob and Nob have the same witch in mind although the speaker does not
commit himself to the claim that witches exist. This is already something of a
poser, to be sure, and the binding theory leads us to expect that the problem
of intentional identity goes even deeper than Geach realized: it should not
be restricted to pronouns but extend to all presupposition-inducing
expressions. This expectation turns out to be correct, as the following
examples illustrate:

(12)    a.    Hob thinks that his cow can whistle, and Nob wonders
              whether his sow can whistle, too.

        b.    Hob thinks that his chickens don't lay eggs anymore, and Nob
              wonders when they'll start laying eggs again.

In the second conjunct of (12a), the presupposition is triggered that some
individual other than Nob's sow can whistle, and apparently this
presupposition can be bound within Hob's belief context in the first
conjunct; for a speaker can utter this sentence without accepting this
presupposition (indeed, as far as the speaker is concerned, Nob may not
even have a sow). The same holds, mutatis mutandis, for (12b), where the
adverbial *again* triggers the presupposition that Hob's chickens have
stopped laying eggs.

The 'intentional identity' cases are more dramatic than the other data
considered in this section, because they suggest that presuppositional
binding may cut across attitude reports on different subjects. This is why the
theory that I propose below does not apply to Hob-Nob sentences. I admit
this unblushingly, because I don't know of any halfway convincing account of
these cases. But it is nonetheless a cause for concern, since I am in two minds
as to whether an analysis of single-subject cases shouldn't apply to Geach's
multiple-subject cases, too. On the one hand, it doesn't seem as if the Hob-
Nob cases are something out of the ordinary, from an observational point of
view; for in their own way, they conform to the familiar patterns of
presuppositional and anaphoric binding:

(13)    a.    If a witch has blighted Bob's mare, then she has killed Cob's
              sow, too.

        b.    If a witch has blighted Bob's mare, then Nob wonders whether
              she has killed Cob's sow, too.

        c.    If Nob believes that a witch has blighted Bob's mare, then he
              (Nob) wonders whether she has killed Cob's sow, too.

> d.  If Hob thinks a witch has blighted Bob's mare, then Nob wonders whether she killed Cob's sow, too.

(14)  a.  Perhaps a witch has blighted Bob's mare, and perhaps she has killed Cob's sow, too.

   b.  Perhaps a witch has blighted Bob's mare, and perhaps Nob wonders whether she has killed Cob's sow, too.

   c.  Perhaps Nob believes that a witch has blighted Bob's mare, and perhaps he (Nob) wonders whether she has killed Cob's sow, too.

   d.  Perhaps Hob thinks a witch has blighted Bob's mare, and perhaps Nob wonders whether she killed Cob's sow, too.

(15)  a.  ?Perhaps a witch has blighted Bob's mare. She has killed Cob's sow, too.

   b.  ?Perhaps a witch has blighted Bob's mare. Nob wonders whether she has killed Cob's sow, too.

   c.  ?Perhaps Nob believes that a witch has blighted Bob's mare. He (Nob) wonders whether she has killed Cob's sow, too.

   d.  ?Perhaps Hob thinks a witch has blighted Bob's mare. Nob wonders whether she killed Cob's sow, too.

In each of (13) through (15), the four variants (a)-(d) pattern together, regardless of whether attitude reports are involved or not, and regardless of whether two consecutive attitude reports are on the same subject, as in the (c) examples, or on different subjects, as in the (d) examples. These observations suggest that, from an observational point of view, there is nothing special about the Hob-Nob connection, but on the other hand I am tempted to concur with Asher's (1987) judgment that there *is* something special about Hob-Nob sentences, after all, and that there is no reason to expect that an adequate analysis of single-subject cases should account for the multiple-subject cases, too. Unfortunately, I will have to leave this matter open. The two theories that will be discussed in the following (Heim's and my own) only apply to the single-subject cases.

## 5.2  No satisfaction

It is not entirely unfair to say that the extant literature on presuppositions yields just a single extensive analysis of the behaviour of presuppositions in attitude contexts, namely Heim's 1992 article.[4] However, the reason why I

---

[4]  For further discussion of the interplay between presuppositional expressions and attitude reports, see Karttunen (1973, 1974), Fauconnier (1985: 93–94, 105–108), van der Sandt (1988: 223–232), Kay (1992), and Zeevat (1992).

have chosen to review Heim's proposal here is not that it happens to lack competition. Rather, it is because in certain respects the theory is very similar to the approach that I want to advocate, while on the other hand there are non-trivial differences as well, and therefore bringing out what is common to the two theories and what is not will give us a sharper view on both. Heim's proposal is a version of what I have called the 'satisfaction theory' of presupposition. In the foregoing chapters I have argued at length and *ad nauseam* that the satisfaction theory is a non-starter, so there is ample reason to reject the Heim's proposal right away. Nonetheless, I will try and judge her account of attitude reports on its own merits.

Given that Heim is working within the framework of the satisfaction theory, it is clear how she must proceed: she will have to devise definitions that capture the truth-conditional meaning of the attitude verbs as well as their projection characteristics. In her article Heim does this for the verbs *believe, want, wish,* and *be glad,* but here we confine our attention to the first two. Heim's analysis of *believe* is a transposition of the classical possible-worlds semantics of belief sentences (Hintikka 1962, 1969). According to the classical analysis, a sentence of the form '*a* believes that $\varphi$' is true in a world $w$ iff $\varphi$ is true in all worlds that are compatible with *a*'s beliefs in $w$. This translates to the context-change framework in a straightforward way, as follows (Heim 1992: 189): Let dox be a function that, in each world $w$, assigns to an individual *a* the set of worlds that are consistent with what *a* believes in $w$, i.e. $dox_a(w)$ is the set of doxastic alternatives that are open to *a* in $w$. Then:

(16)     $c[a \text{ believes } \varphi] = \{w \in c \mid dox_a(w) \text{ satisfies } \varphi\}$

As before, a context c satisfies $\varphi$ iff $c[\varphi] = c$. Since contexts and sets of doxastic alternatives are of the same type (i.e. they both are sets of worlds), it makes sense to regard the latter as contexts, too, on which context-change functions operate in the usual way. Accordingly, we shall refer to sets of doxastic alternatives as 'doxastic contexts'.

What (16) does is eliminate from the current context set all those worlds in which *a* does not believe that $\varphi$. With respect to the projection properties of *believe,* (16) predicts that '*a* believes that $\varphi\{\chi\}$' presupposes that *a* believes $\chi$: $c[a \text{ believes } \varphi\{\chi\}]$ is guaranteed to be defined only if in all worlds $w \in c$, $dox_a(w)$ satisfies $\chi$, which means that c must already satisfy that *a* believes $\chi$. For example, the *it*-cleft in (17a) triggers the presupposition that (17b) is true, and when the *it*-cleft is embedded in a belief sentence, as in (18a), Heim's theory predicts that the resulting presupposition is (18b):

(17)     a.     It was Fred who tripped Barney.

          b.     Barney was tripped.

(18)     a.     Barney believes that it was Fred who tripped him.

         b.     Barney believes that he was tripped.

Heim's theory correctly predicts that, whereas (19a) will not give rise to the inference that (18b) is true, (19b) will:

(19)     a.     Barney believes that he was tripped, and he believes that it was Fred who tripped him.

         b.     Perhaps Barney believes that it was Fred who tripped him.

In (19a) the presupposition that (18b) is true is satisfied by the first conjunct, and therefore the sentence as a whole entails but does not presuppose that Barney believes that he was tripped. On the other hand, if (19b) is uttered in a context in which it is not already given that (18b) is true, the definition in (16) requires that this information be accommodated (I assume here that Heim can somehow account for the fact that modals are holes). Both predictions are correct, at least in the sense that the inferences that the theory predicts accord with speakers' intuitions.

    On the negative side, although Heim's definition correctly predicts that the presupposition that (18b) is true is suspended in (19a), it also predicts that (20a) presupposes (20b):

(20)     a.     It is possible that Barney was tripped, and that he believes that it was Fred who tripped him.

         b.     Barney was tripped → Barney believes that he was tripped

This prediction is clearly false, for (20a) does not imply (20b) in any way. (Incidentally, this counterexample does not hinge upon Heim's analysis of conjunctions. If that were the case, we should expect her predictions to be too weak, not incorrect.) This example indicates that presuppositions that originate in belief contexts may sometimes be bound, or satisfied, in a non-belief context, and the problem with Heim's theory is that it doesn't allow for this.

    The second problem is related to the first one. It is that Heim's proposal only accounts for one half of the two-sided interpretations that presuppositions in attitude contexts often give rise to. (16) does not allow us to infer from (19b) that, according to the *speaker*, Barney was tripped, although intuitively an utterance of (19a) would normally license this inference. But as Heim agrees with Karttunen that (18b) is more basic in that it, rather than (17b), should be accounted for in presuppositional terms (i.e. she accepts the i-principle rather than the e-principle), she still has to explain how the inference to (17b) comes about. Later on in this section we shall see how she proposes to do this, but before we turn to that issue, let us first have a look at Heim's semantics of *want*.

For the verb *want*, Heim considers two possible analyses. One of these mirrors her treatment of *believe*. Besides a function that assigns subjects doxastic contexts, we assume as given a function whose range consists of sets of buletic alternatives: in each world $w$, $\text{bul}_a(w)$ is the set of worlds that are compatible with $a$'s wishes and desires in $w$. The context-change semantics of *want* is then obtained by substituting '$\text{bul}_a(w)$' for '$\text{dox}_a(w)$' (Heim 1992: 192):

(21)     $c[a \text{ wants } \varphi] = \{w \in c \mid \text{bul}_a(w) \text{ satisfies } \varphi\}$

Not surprisingly, the predictions that this yields parallel those of (16) to a T.

(22)     a.     Harry wants to have a son and he wants his son to be the first pianist who can play the Moonlight Sonata in less than six minutes.

           b.     Harry wants his son to be the first pianist who can play the Moonlight Sonata in less than six minutes.

           c.     Harry wants to have a son.

           d.     Harry has a son.

The definition in (21) predicts that the second conjunct of (22a) presupposes (22c), but that this presupposition is absorbed by the first conjunct. In (22b), in contrast, the same presupposition is not neutralized, and therefore it is predicted that this sentence presupposes that Harry wants to have a son. Both predictions are obviously false: we would not normally infer from (22b) that (22c) is true, nor would we want to say, intuitively speaking, that this inference is neutralized in the process of interpreting (22a). In a word: '$a$ wants $\varphi\{\chi\}$' does not normally imply, let alone presuppose, that $a$ wants $\chi$.

As far as its predictions about the presuppositional properties of *want* are concerned, the proposal in (21) has nothing whatsoever to recommend it, and this would be the end of the story if it were the case that (21) adequately represented the meaning of *want*. However, Heim argues that this rule is already unsatisfactory for truth-conditional reasons alone. In particular, it follows from (21) that if $a$ wants $\varphi$ then $a$ must want all logical consequences of $\varphi$. But, for example, the desire to start a relief fund for orphaned sheep does not entail the desire that there be orphaned sheep. Therefore, Heim proposes an alternative semantics for *want*, which is inspired by Stalnaker's characterization of wanting:

           [...] wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he wants. Some propositions which are entailed by propositions that one

wants to be true in this sense are also entailed by the relevant alternatives. It is not that I want these propositions to be true — it is just that I accept that they will be true whether I get what I want or not. (Stalnaker 1984: 89)

Stalnaker suggests that '*a* wants φ' must be evaluated against the background of a set of 'relevant alternatives'. What alternatives (i.e. possible worlds) are these? Stalnaker is not overly explicit on this point, but it seems plausible to assume, and in conformity with what he says elsewhere, that the set of relevant alternatives at any world $w$ is the set of doxastic alternatives open to $a$ in $w$, i.e. $\text{dox}_a(w)$. Stalnaker's suggestion may then be explicated as follows:

(23)     '*a* wants φ' is true in a world $w$ iff $a$ prefers all φ-worlds in $\text{dox}_a(w)$ to any non-worlds in $\text{dox}_a(w)$.

This smoothly translates into the context-change function for *want* as given in (24). Let '$<_{a, w}$' denote an ordering which ranks sets of worlds according to their desirability for $a$ in $w$. So if $p$ and $q$ are sets of worlds, '$p <_{a, w} q$' is intended to mean that, in $w$, $a$ prefers all worlds in $p$ to any world in $q$. Now the context-change function associated with '*a* wants φ' may be defined as follows:

(24)     $c[a \text{ wants } \varphi] = \{w \in c \mid \text{dox}_a(w)[\varphi] <_{a, w} \text{dox}_a(w)[\neg\varphi]\}$

This is not quite the same as the rule proposed by Heim, which is more involved, but since the complications that Heim argues for are not relevant to our purposes, we can make do with a simplified version. Needless to say, none of the critical remarks that I shall make is contingent up upon this simplification.

The meaning rule given in (24) predicts the following presuppositional profile for *want*: (24) entails that '*a* wants φ{χ}' is defined in a context c iff φ{χ} is defined in $\text{dox}_a(w)$, for all $w \in c$, which means that '*a* wants φ{χ}' presupposes that $a$ believes χ. Thus it is predicted, correctly, that we would normally infer from (22b), but not from (25), that Harry believes that he has a son.

(25)      Harry believes that he has a son and he wants his son to be the first pianist who can play the Moonlight Sonata in less than six minutes.

So the revised semantics of *want* validates the i-principle, which is one of Heim's avowed aims. It therefore does slightly better than its predecessor, but it is not exactly a great leap forward. One problem with (24), which Heim acknowledges (1992: 200), is that it doesn't explain presupposition filtering in *want-want* sequences: (22a) is predicted to imply that Harry believes that

he has a son, which is clearly false. Also, (26a) is now predicted to presuppose (26b), which is wrong, too:

(26)    a.    If Harry wants to have a son, he wants his son to be the first pianist who can play the Moonlight Sonata in less than six minutes.

        b.    Harry wants to have a son $\rightarrow$ Harry believes that he has a son

One of the objectives of Heim's article is to derive the i-principle, and in this particular respect, at least, Heim's proposals are successful. It remains to be seen, however, how the resulting theory can be reconciled with the observation that presuppositional expressions in attitude contexts will often receive a two-sided reading. Supposing for argument's sake that Heim's theory of presupposition is correct, her problem is to account for the fact that, in some cases at least, a presuppositional expression $\chi$ which occurs in an attitude context not only gives rise to the inference that the subject of the attitude believes that $\chi$, but also licenses the inference that the speaker holds $\chi$ to be true. Heim considers two possible solutions to this problem.

The first possible solution that Heim suggests is that, in those cases in which a presupposition appears to escape from an attitude context, the expression that has triggered it is being given a *de re* interpretation. To illustrate, Heim's suggestion is that, possibly, we infer from (27a) that (27b) is true because *Müller's Requiem* is construed as having scope over the attitude verb:

(27)    a.    The Osnabrück Philharmonic wants to put on Müller's Requiem.

        b.    Müller has composed a Requiem.

Heim reasons as follows. The satisfaction theory does not license the prediction that (27a) presupposes (27b). But it is well known that it is often possible to construe noun phrases *de re*. So *instead of* explaining the inference from (27a) to (27b) in presuppositional terms, it might be viewed as the result of a *de re* construal, carried out by a mechanism that is motivated on independent grounds.

This argument runs into several difficulties, some of which Heim notes herself. First, it is not at all obvious that other presuppositional constructions besides definite noun phrases can be read *de re*. Heim tries to show that this is possible for aspectual verbs and focus particles like *also*. (She doesn't consider any further presupposition inducers, so even if her analysis of these two is correct, which I doubt, a wide range of expressions remain to be considered.) An example she discusses is (28a), which she claims may be read, *à la* Kaplan (1969), as in (28b):

(28)    a.    John thought I had stopped proof-reading.

   b.    There is an acquaintance relation D such that (i) John bore D
         to my proof-reading, and (ii) John thought that the activity he
         bore D to had stopped. (Heim 1992: 208)

In other words, Heim claims that, conceivably, (28a) may mean that '[...]
John thought of the activity of mine that was in fact a proof-reading, but that
he may not have recognized as such, that it had stopped.' (Heim 1992: 208)
Let us grant that a scenario can be constructed that enforces such a reading.
According to Heim, this reading can be obtained by construing *proof-reading*, in effect, as a wide-scope definite. But this cannot be right. Consider
the following minimal variant of (28a), for instance:

(29)    John thought I had started proof-reading.

If Heim's analysis of (28a) were correct, then it should be possible to read
(29) as meaning that John thought of the activity of mine that was in fact a
proof-reading, but that he may not have recognized as such, that it had
started. If this makes sense at all, it is not the reading that we want to account
for. What needs to be explained, rather, is how (29) can, and typically will,
give rise to the inference that the speaker *wasn't* proof-reading before the
reference time, and it is by no means clear that this can be done by means of
*de re* construal.

   In § 1.4 it was shown that scope taking and presupposition projection are
quite different things. If we say that an expression $\alpha$ takes wide scope, for
example, then we are speaking of $\alpha$ itself or a unit corresponding with $\alpha$ at
some level of analysis (such as $\alpha$'s counterpart at LF, or its denotation with
respect to a given model, or whatever). If, on the other hand, we say that $\alpha$
triggers the presupposition that $\chi$, then $\chi$ is a requirement that $\alpha$ imposes on
its context. So, whereas presupposition projection clearly is a pragmatic
affair, scope taking is primarily a grammatical phenomenon. At the end of
the day, it is this distinction which rules out the possibility of explaining
presupposition in terms of scope. But if that is impossible, then the same
applies for Heim's proposal that external readings of presuppositional
expressions might be accounted  for in terms of *de re* construal. For even if
there is a difference between construing an expression *de re* and giving it
wide scope, the former surely involves something very much like the latter,
and no matter how exactly *de re* construals are to be accounted for, all the
problems discussed in Chapter 1 are problems for Heim's proposal, too.

   Heim assumes that there are two mechanisms, so to speak, which can make
it seem as if a presuppositional expression has escaped from a local context.
On the one hand there is a presupposition-projection mechanism, which she
claims is derivable from a semantics in terms of context change, and on the
other hand there is the mechanism for generating *de re* construals. But
although it can hardly be denied that two such mechanisms exist, it is not at

all obvious that they both apply to presuppositional expressions. If presuppositions already have a natural tendency to float up from embedded positions anyway, why should they need help from a mechanism for *de re* construal? I take it that we should prefer a theory that is in a position to say that *de re* interpretations of presuppositional expressions are the outcome of the standard mechanisms of presupposition projection, and that a special mechanism for *de re* construal is only required in exceptional cases, as with marked indefinite noun phrases. Given the theory that I present below, all typical *de re* construals fall out automatically as instances of presupposition projection, and I take it that, ceteris paribus, this type of account is to be preferred to Heim's.[5]

One final remark about Heim's calling upon a mechanism for *de re* construal: it seems to me that in doing so she is pulling the rug from under her own theory. If it is indeed the case that presuppositional expressions in general can be construed as taking scope over the immediate contexts in which they occur, it is unclear why an appeal to wide-scope readings should be restricted to attitude contexts. Why not describe *all* presuppositional phenomena in terms of scope? For reasons discussed in § 1.4 this is not a realistic option. It should be clear however that the manoeuvre which Heim is contemplating is liable to backfire on her own theory.

To recapitulate: in the first half of this section I have argued that the semantics which Heim provides for the verbs *believe* and *want* fails to correctly predict the presuppositional properties of these verbs. This is not to say that this semantics is wrong, for in fact I intend to adopt it myself, too. But it is to say that, as in the case of the connectives, the projection profile of an attitude verb does not follow from its meaning alone. In the second half of this section I partly reviewed Heim's attempts to explain the fact that presuppositions triggered in attitude contexts may receive a two-sided interpretation. Since Heim's theory only accounts for the internal component, she has to explain the two-sided reading on the basis of an internal construal. I have argued against one of the two lines of argument that Heim suggests for this job. Another line she tentatively pursues will be discussed in § 5.6.

### 5.3 Belief in DRT

My account of presuppositions in attitude contexts is similar to Heim's in a number of ways, but crucially different in others, the most significant

---

[5] It will have become clear in the meantime what led me to introduce the distinction between internal and external construals of presuppositional expressions in addition to the *de re/de dicto* distinction.

difference being, of course, that I adopt the binding theory. Another important difference is that, contrary to Heim, I take the e-principle as my starting point. This is not to say that I repudiate the i-principle altogether. What I take to be correct about this principle is the notion that belief is the central case in the sense that a presupposition $\chi$ that is triggered within an attitude report on person $a$ can be satisfied, or bound, by what is being assumed about $a$'s beliefs. But unlike Heim or Karttunen, I don't take this to imply that the essential datum to be explained is how such a presupposition can end up as a presupposition of the form '$a$ believes that $\chi$'. The main question is rather why attitude verbs are generally transparent to presuppositions, and in this respect at least my perspective is exactly the opposite from Heim's.

The theory of beliefs and wants presented below is the same as Heim's semantically speaking: like Heim, I adopt a possible-worlds semantics for belief reports, and concerning the semantics of *want* I will take Stalnaker's analysis as my point of departure, following Heim in this respect, too. But since I am not committed to the satisfaction theory, I do not have to claim that the projection characteristics of *believe* and *want* follow from their meanings alone. In fact, what I shall be claiming is that the projection profile of attitude verbs generally is determined by their meanings *and* by the presuppositions that they induce.

In this section and the next one we will be concerned with the representation in a DRT framework of beliefs and belief reports; presuppositions will not re-enter the stage until § 5.5. Prior DRT accounts of belief and belief reports have been presented by Zeevat (1984, 1989b), Asher (1986, 1987, 1993), and Kamp (1987, 1990), who propose a structural theory of belief (the term is Kamp's), which is to say that belief is construed as a relation between an individual and a syntactic object, i.e. a DRS. The theory that I advocate is not a structural one. Although DRT is a natural framework for articulating a structural theory of belief, the two are not tied together. One can adopt DRT as a framework for analysing belief *reports* and at the same time construe belief as a relation between individuals and sets of worlds, which is what I propose to do.[6]

The classical possible-worlds analysis of belief (and the other attitudes) is plagued by a family of well-known puzzles having to do with logical omniscience and (non-) substitutivity of synonymous expressions, and one of the main arguments in favour of the structural account is that it manages to avoid these problems, because its principal tenet is that the objects of belief are entities with a rich internal structure. If I hang onto the possible-worlds analysis nonetheless it is because I feel that it is more consonant with the

---

[6]   Asher (1986) develops such a theory, too, but does not actually endorse it, inter alia for reasons discussed in the next paragraph.

common sense notion of belief than any other account that I know of, and I have not given up hope yet that its problems can be solved, for instance along the lines set out by Stalnaker (1984); but I will not address these problems here.

So let us assume that belief is a relation between an individual and a set of worlds. More accurately, if someone asserts that '*a* believes φ', he claims that φ is satisfied by *a*'s doxastic context, i.e. the set of worlds that are compatible with *a*'s beliefs. If we want to embed this analysis into the DRT framework, we will need some device for representing doxastic contexts. To this purpose, I introduce what I call 'propositional terms', which denote relations between worlds and embedding functions. Or, what amounts to the same thing, propositional terms denote sets of pairs consisting of a world and an embedding function. Such pairs may be viewed as partialized worlds, with the embedding function labelling a limited number of objects at a given world. I will use the term 'indexed worlds' to refer to pairs of worlds and embedding functions; sets of indexed worlds will be called 'indexed propositions'. Each indexed proposition σ uniquely determines a set of worlds, $w(\sigma)$, which I will sometimes refer to as the 'classical proposition' corresponding to σ:

(30)    $w(\sigma) = \{w \mid \exists f: \langle w, f \rangle \in \sigma\}$

I propose to extend the original DRT language in an essentially conservative way. The main difference between the following version of DRT and the standard one is that the latter has a poorer inventory of terms.

(31)    *DRS-terms*

    a.   $T = Rm_i \cup RM_p \cup T_p$ (terms)

    b.   $RM_i = \{u, v, w, ..., u', u'', ...\}$ (individual reference markers)

    c.   $RM_p = \{p, q, r, ..., p', p'', ...\}$ (propositional reference markers)

    d.   $T_p = \{p+\varphi \mid p \in RM_p \text{ and } \varphi \text{ is a DRS}\}$ (complex propositional terms)

Whereas in the original version of DRT there are only individual reference markers, the present version also features propositional terms (i.e. members of $RM_p \cup T_p$). The denotations of propositional terms will be indexed propositions, and the denotation of a complex propositional term $p+\varphi$ will be the indexed proposition denoted by $p$ incremented with the information in φ.

(32)    *DRSs and DRS-conditions*

    a.   A DRS φ is a pair $\langle U(\varphi), Con(\varphi) \rangle$, where $U(\varphi) \subseteq RM_i \cup RM_p$, and $Con(\varphi)$ is a set of DRS-conditions.

   b.    If P is an *n*-place predicate, and $\alpha_1, \ldots, \alpha_n \in T$, then $P\alpha_1 \ldots \alpha_n$
         is a DRS-condition.

   c.    If $\alpha \in RM_i \cup RM_p$ and $\beta \in T$, then $\alpha = \beta$ is a DRS-condition.

   d.    If $\varphi$ and $\psi$ are DRSs, then $\neg\varphi$, $\varphi \Rightarrow \psi$, and $\varphi \vee \psi$ are
         DRS-conditions.

Note that (32c) only caters for identity conditions in which a reference
marker occurs on the left-hand side. Expressions of the form '$q = p+\varphi$'
(where $p$ and $q$ are propositional reference markers) are therefore
admissible conditions; expressions of the form '$p+\varphi = q+\psi$' or '$p+\varphi = q$' are
not. The reason for this is purely technical. On the one hand, we won't be
needing any other identity conditions than those admitted by (32c). On the
other hand, if we allowed the left-hand slot of an identity condition to be
occupied by any term, the definition of DRS-accessibility would become
more complex.

The following example illustrates how belief reports may be represented
in the language defined in (31)-(32) (in this section we are concerned with
representations like these and what they mean; later it will be shown how
they are constructed):

(33)    a.    There is an A who believes that there is a B.
        b.    [x, p, q: Ax, x believes p, x believes q, q = p+[y: By]]

Where, as one might expect, *believes* is a two-place predicate linking
individual and propositional terms. That, on the intended interpretation
sketched above, (33b) is an adequate representation of the content of (33a)
may be seen as follows. The propositional reference markers p and q are
partial representations of x's beliefs, where q increments p with the
information that there is a y such that By. Suppose that x stands in the
*believes* relation to p, i.e. p is a correct but possibly incomplete picture of x's
doxastic context. Then, if (33a) is true, x should stand in the *believes* relation
to q, too. (There are other ways of representing belief reports with the help
of the language defined by (31)-(32). In § 5.5 it will become clear why I have
chosen this particular format.)

A minor inconvenience caused by representations like (33b) is that they
tend to be cluttered by sequences of conditions of the form '*u* believes
$p_1, \ldots, u$ believes $p_n$'. To alleviate this problem, I will abbreviate such
sequences as '*u* believes $p_1$ & $\ldots$ & $p_n$'. So the following is a shorthand for
(33b):

(33)    c.    [x, p, q: Ax, x believes p & q, q = p+[y: By]]

If a speaker issues several belief reports on the same subject *a*, he may be
construed as presenting a sequence of increasingly informative pictures of *a*'s
beliefs.

(34)    a.    There is an $A_x$ who believes that there is a $B_y$, and he$_x$ also believes that she$_y$ is a C.

        b.    $[_1$ x, p, q, r: Ax, x believes p & q & r,
              q = p+$[_2$ y: By], r = q+$[_3$ : Cy]]

The indices in (34a) serve to indicate co-reference, as usual, but furthermore the index on a noun phrase is a homograph of the reference marker which represents that noun phrase in (34b); this convention (which is *not* part of the theory) is meant to clarify the links between between NPs and their semantic representations. In (34b), q is an extension of p, and q is extended by r: q adds to p the information that there is a y such that By, and r adds to q the information that Cy. And since x stands in the *believes* relation to p, q, and r, he must believe that there is a y such that By and Cy.

(35)    a.    If an $A_x$ believes that there is a $B_y$, he$_x$ also believes that she$_y$ is a C.

        b.    $[_1 : [_2$ x, p, q: Ax, x believes p & q, q = p+$[_3$ y: By]] $\Rightarrow$

              $[_4$ r: x believes r, r = q+$[_5$ : Cy]]]]

The DRS in (35b) represents the content of (35a), and its intuitive interpretation is the following. Let x be an arbitrary object such that Ax, and suppose that x's doxastic context is correctly characterized by some p and q, where q extends p with the information that there is a y such that By: then it must be the case that there is an r which extends q with the information that Cy such that x's doxastic context is correctly characterized by r.

Examples like (34) and (35) show that in the present version of DRT, a DRS may contain identity conditions that affect the extension of the accessibility relation. (34b$_2$) is accessible to (34b$_3$) only because the former extends a propositional term, p, yielding a term q which in its turn is extended by the latter; the same holds for (35b$_3$) and (35b$_5$). However, in neither case do we want to say that the main DRS is accessible to these embedded DRSs, for the following reason. Saying that a DRS φ is accessible to another DRS ψ is tantamount to saying that the context represented by ψ is an extension of the context represented by φ, at least in the sense that the denizens of φ inhabit ψ, too. In its entirety, (35b) represents the commitment slate of given speaker, while (35b$_3$) and (35b$_5$) represent the doxastic context of somebody else, viz. an arbitrary A. So we shouldn't even require that the information in the DRSs be consistent with the information contained in the main DRS. Hence, (35b$_1$) is not accessible to (35b$_3$) or (35b$_5$); analogous remarks apply to (34b).

These observations prompt the following definitions. I define DRS-accessibility in tandem with an 'extension' relation between propositional terms, which is given in (36). The relation defined in (37) is called '*strict*

accessibility' because a second, less restrictive, notion of accessibility will be introduced later on.[7]

(36)    *Extension*

For any $\varphi$, $\leq_\varphi$ is the smallest preorder on propositional terms for which the following hold, for all $\psi \leq \varphi$, and $p, q \in RM_p \cup T_p$:

a.    If $p = q{+}\chi \in Con(\psi)$, then $q \leq_\varphi p$

b.    If $p = q \in Con(\psi)$, then $p \leq_\varphi q$ and $q \leq_\varphi p$

(37)    *Strict accessibility*

$\leq$ is the smallest preorder for which the following hold:

a.    If $\neg\psi \in Con(\varphi)$, then $\varphi \leq \psi$

b.    If $\psi \vee \chi \in Con(\varphi)$, then $\varphi \leq \psi$ and $\varphi \leq \chi$

c.    If $\psi \Rightarrow \chi \in Con(\varphi)$, then $\varphi \leq \psi \leq \chi$

d.    If $p{+}\psi \leq_\varphi q{+}\chi$, then $\psi \leq \chi$

(38)    *Accessible domains*

The accessible domain of a DRS $\varphi$ is the set of reference markers that are 'visible' from $\varphi$:

$Acc(\varphi) = \cup\{U(\psi) \mid \psi \leq \varphi\}$

$\leq_\varphi$ and $\leq$ may be viewed as information orderings. If $\varphi \leq \psi$, then $\psi$ contains at least as much information as $\varphi$ does. Likewise, if $p \leq_\varphi q$, then $q$ contains at least as much information as $p$.

To illustrate the interdependence between $\leq_\varphi$ and $\leq$, consider the DRS in (35b) again. It follows from (37c) that $(35b_1) \leq (35b_2) \leq (35b_4)$, and then it follows from (36a) that q $\leq_{(35b4)}$ r, and since q = p+$(35b_3)$ and r = q+$(35b_5)$, (37d) gives us that $(35b_3) \leq (35b_5)$.

The notion of strict accessibility is defined in such a way that, for example, $(35b_1)$ is accessible neither to $(35b_3)$ nor to $(35b_5)$ and so these sub-DRSs don't have access to the reference markers introduced in the main DRS: x, p, q, and r. Consequently, (39a) is not a proper DRS:[8]

---

[7]    Wanting to avoid an orgy of subscripts and superscripts, I decided to leave these definitions incomplete in two respects. First, it is understood that (36) and (37) are about expression tokens, not types. Secondly, $\leq$ and $\leq_\varphi$ actually are three- and four-place relations, respectively, since both are only defined within the context of a larger DRS, strictly speaking.

[8]    As usual, a proper DRS is one in which all reference markers are 'bound', i.e. $\varphi$ is a proper DRS iff none of its sub-DRSs $\psi$ contain a condition employing a reference marker $u$ such that $u \notin Acc(\psi)$.

(39)    a.    [x, y, p, q: Ax, x believes p & q, q = p+[: By]]

        b.    [x, y, p, q: Ax, x believes p & q, q = p+[y: By]]

In (39a), the DRS [: By] employs a reference marker, y, which is not in its accessible domain (although it occurs in the universe of the main DRS), and therefore (39a) is not a proper DRS. By the same token, in (39b), which *is* a proper DRS, the reference marker y is *re*-introduced in the embedded DRS, or to put it otherwise, the occurrence of y in the embedded DRS is not 'bound' in the main DRS. Hence, the fact that the same reference marker occurs in the main DRS and a belief DRS is meaningless in the system as it stands. Later on, we will exploit this circumstance and endow such pairs of reference markers with a useful interpretation.

Let us now turn to the model-theoretic interpretation of the language defined above. The semantic values of DRSs and propositional terms will be of the same type, as they both denote what I have called 'indexed propositions': sets of indexed worlds $\langle w, f \rangle$, where $w$ is a world and $f$ is an embedding function. Given that an embedding function is to assign propositional reference markers values that contain embedding functions, it will be clear that defining the notion of embedding function is a somewhat delicate undertaking, because the most straightforward definition would lead to paradox.

(40)    *Embedding functions*

        Let M be a model with a set of worlds W and a domain of individuals $D_w$ for each $w \in W$. Then:

        a.    $F_M^0 = \{f \mid \exists X \subseteq RM_i \text{ such that } f : X \to \cup\{D_w \mid w \in W\}\}$

        b.    $F_M^n = \{f \cup g \mid f \in F_M^0 \text{ and } \exists X \subseteq RM_p \text{ such that } g : X \to Pow(W \times F_M^{n-1})\}$, if $n > 0$

A 0-order embedding function simply is a partial function from individual reference markers to individuals. For all $n > 0$, each $n$-order embedding function consists of two components, both of which may be empty: (i) a 0-order embedding function and (ii) a partial function from propositional reference markers to sets of indexed worlds $\langle w, f \rangle$, where $w$ is a world and $f$ is an $n$-1-order embedding function. Note that the resulting hierarchy of embedding functions is an inclusive one in the sense that, for all $n \geq 0$, $F_M^n \subseteq F_M^{n+1}$

Here is a picture of an embedding function (a comparatively simple one):

(41)
$$\begin{bmatrix} x \to a \\ y \to b \\ p \to \left\{ \begin{array}{c} \langle w_1, g{:}[x \to c] \rangle \\ \langle w_2, h{:}[y \to d] \rangle \end{array} \right\} \end{bmatrix}$$

$f$ is an embedding function that maps the individual reference markers x and y onto $a$ and $b$, respectively; the propositional reference marker p is mapped onto an indexed proposition consisting of two indexed worlds, which contain two further embedding functions, $g$ and $h$.

The models for our DRS language come with a function called 'dox', which assigns belief contexts to individuals in worlds, as in Heim's theory. For any individual $a$ and world $w$, $\mathrm{dox}_a(w)$, if defined, is a set of worlds. The interpretation of the two relational constants *believes* and *considers* is defined in terms of this function, as follows:

(42)    a.    $I_w(\text{believes}) = \{\langle a, \sigma \rangle \mid \mathrm{dox}_a(w) \subseteq w(\sigma)\}$

        b.    $I_w(\text{considers}) = \{\langle a, \sigma \rangle \mid \mathrm{dox}_a(w) \cap w(\sigma) \neq \varnothing\}$

The semantic correlates of DRSs and propositional terms are indexed propositions, and each indexed proposition $\sigma$ uniquely determines a classical proposition $w(\sigma)$. Thus (42a) says that $a$ bears the *believes* relation to $\sigma$ in a world $w$ iff $a$'s doxastic context entails the classical proposition determined by $\sigma$. (42b) says that $a$ bears the *considers* relation to $\sigma$ in a world $w$ iff $a$'s doxastic context is compatible with $w(\sigma)$. So, *considers* stands to *believes* as '$\exists$' stands to '$\forall$'.[9]

(43)    *DRS semantics*

        Let $M = \langle W, D, \mathrm{dox}, I \rangle$ be a model, where W is a set of worlds, D is a function that assigns a domain of individuals $D_w$ to each $w \in W$, dox a partial function from $W \times D$ to $\mathrm{Pow}(W)$ and I an interpretation function. Let $s = \langle w, f \rangle$ be an indexed world, where $w \in W$ and $f \in F_M^n$, for some $n \geq 0$. Then:

        a.    $[\![\alpha]\!]_s = f(\alpha)$, if $\alpha \in RM_i \cup RM_p$ ; undefined otherwise

        b.    $[\![p{+}\varphi]\!]_s = \{s' \mid \exists s'' \in [\![p]\!]_s : s' \in [\![\varphi]\!]_{s''}\}$

---

[9]  Perhaps I should note explicitly that *believes* and *considers* are defined here as technical terms, and that it is not my intention to claim that these terms, taken on their own, capture the intuitive meaning of the verbs of which they happen to be homographs. This holds in particular for the second predicate, whose English gloss will not always be 'considers' but rather something like: 'does not believe that not'.

c.  $[\![\varphi]\!]_s = \{\langle w, g\rangle \mid f \subseteq g$ and $\mathrm{dom}(g) = \mathrm{Acc}(\varphi)$ and
$\forall\psi \in \mathrm{Con}(\varphi)\colon [\![\psi]\!]_{\langle w, g\rangle} = 1\}$

d.  $[\![P\alpha_1 \ldots \alpha_n]\!]_s = 1$ iff $\langle[\![\alpha_1]\!]_s, \ldots, [\![\alpha_n]\!]_s\rangle \in I_w(P)$

e.  $[\![\alpha = \beta]\!]_s = 1$ iff $[\![\alpha]\!]_s = [\![\beta]\!]_s$

f.  $[\![\neg\varphi]\!]_s = 1$ iff $[\![\varphi]\!]_s = \varnothing$

g.  $[\![\varphi \vee \psi]\!]_s = 1$ iff $[\![\varphi]\!]_s \cup [\![\psi]\!]_s \neq \varnothing$

h.  $[\![\varphi \Rightarrow \psi]\!]_s = 1$ iff $\forall s' \in [\![\varphi]\!]_s\colon [\![\psi]\!]_{s'} \neq \varnothing$

In order to illustrate how this semantics works, I give the interpretation of (39b):

(44)  $[\![(39b)]\!]_{\langle w, f\rangle} =$

$\{\langle w, g\rangle\mid f \subseteq g$ & $\mathrm{dom}(g) = \{x, p, q\}$ & $g(x) \in I_w(A)$ &

$\mathrm{dox}_{g(x)}(w) \subseteq w(g(p)) \cap w(g(q))$ &

$g(q) = \{\langle v, j\rangle \mid \langle v, h\rangle \in g(p)$ & $h \subseteq j$ & $\mathrm{dom}(j) = \{y\}$ & $j(y) \in I_v(B)\}\}$

This says, first, that for each indexed world $\langle w, g\rangle$ in the denotation of (39b), $g$ must have $\{x, p, q\}$ as its domain and $g(x)$ must be an A. Secondly, the classical propositions determined by $g(p)$ and $g(q)$ have to be entailed by $\mathrm{dox}_{g(x)}(w)$, which is to say that $g(p)$ and $g(q)$ must provide correct descriptions of $g(x)$'s beliefs. Thirdly, in every indexed world in $g(q)$ there must be a y with property B, so by transitivity $\mathrm{dox}_{g(x)}(w)$ must entail that there is a y with property B. This captures the intended interpretation of (39b).

## 5.4 Counterparts

As it stands, our DRT analysis is restricted to belief reports that are construed strictly *de dicto*; *de re* interpretations cannot be adequately represented.

(45)  a.  Somebody$_x$ believes of something$_y$ that it$_y$ is a sheep.

b.  [x, y, p, q: x believes p & q, q = p+[y: sheep y]]

(45b) is the closest we can get to the most obvious interpretation of (45a), but this is not close enough, because there is no connection between the occurrences of y in the main DRS and in the embedded DRS. That is to say, there is no connection *yet*, because I propose to refine the analysis in such a way that, whenever a constellation like (45b) arises, the admissible values of the two occurrences of y must be counterparts. Hence, the meaning of (45b)

will be roughly the following: 'There is an *a* and a *b*; *a* believes that there is a *c* such that *c* is a sheep; and *b* and *c* are counterparts'.

As Lewis (1968: 114) says, 'The counterpart relation is our substitute for identity between things in different worlds.' But it is a powerful substitute. It allows us to say, for instance, that a thing in one world has several counterparts in another world; or that a thing in one world has no counterpart in another. These are precisely the kind of things that we feel like saying when describing certain intersubjective liaisons between doxastic contexts. As far as Fred is concerned, the prime minister and the president of France are one and the same person. Barney knows better: two jobs, two individuals. Barney has two people in mind who correspond with a single person in Fred's picture of the world: they are counterparts.

According to Lewis, the counterpart relation is a relation of similarity, and if we assume that it may be any kind of similarity relation, this implies that in general the counterpart relation is neither symmetric nor transitive (see Lewis 1968: 115–117). However, it is fairly clear that, certainly in the case of the attitudes but presumably in other cases as well, the counterpart relation cannot be any kind of similarity relation. Barney will say that his notion of Helmut Kohl and the one entertained by Fred are counterparts, because he assumes that Kohl *à la* Fred goes back to the same source as his own Kohl *à la* Barney (cf. Edelberg 1992). It is this type of relation that we typically have in mind when we consider correspondences between belief states, and *this* type of relation is presumably symmetric.

Unlike identity, counterpart relations are indeterminate in various ways.[10] To say that two individuals are counterparts is to say that they are alike in some respects, and since similarity is a matter of degree, we have to agree on a lower bound of similarity before we can decide whether a counterpart relation obtains in any given case. Furthermore, similarity is a perspectival notion. Two individuals may be similar in some respects but not in others, and before we can say that they are counterparts or not we have to decide what respects are to count.

The convention for encoding counterpart relations that I propose goes as follows: if a DRS $\varphi$ is embedded in a DRS $\psi$, and $\varphi$ and $\psi$ share a common reference marker $u$, then the semantic values of $u$ in $\varphi$ and in $\psi$ must be counterparts. In order to make this idea a bit more precise, let us consider the following example:

(46)    a.    Somebody$_x$ believes of someone$_y$ that she$_y$ believes of something$_z$ that it$_z$ is a sheep.

        b.    $[_1$ x, y, p, p': x believes p & p',
              p'= p+$[_2$ y, z, q, q': y believes q & q', q' = q+$[_3$ z: sheep z]]]

---

[10] For more extensive discussion of this point, see Lewis (1983b).

(46b) involves two belief contexts, represented by $(46b_2)$ and $(46b_3)$ one of which is embedded in the other. These two DRSs share a reference marker, i.e. z, and therefore the values of z have to be counterparts. That is, the correlate of z in somebody$_x$'s belief context must be a counterpart to the correlate of z in someone$_y$'s belief context. Furthermore, since $(46b_1)$ and $(46b_2)$ share a reference marker, too, i.e. y, the correlate of y in somebody$_x$'s belief context will have to be a counterpart to the correlate of y in the speaker's commitment slate, which is represented by (46b).

If this is a treatment of *de re* readings, my notion of *de re* construal is a rather weak one. In the case of (46a), for example, the analysis I propose merely says that there is some correlation between what the speaker is committed to and what somebody$_x$ is claimed to believe. Many philosophers would not want to call this a *de re* construal, because they hold, for example, that one can only have *de re* beliefs about objects that one has had intimate cognitive intercourse with. However, save for terminological differences, my account is compatible with virtually any theory of *de re* attitudes that I know of. For my current purposes, the philosophical controversies over *de re* attitudes and attitude reports are largely irrelevant, because my main concern is *that* there is sometimes a correlation between the speaker's commitments and the beliefs he ascribes to someone else; I am less interested in what exactly the nature of such correlations may be, and therefore I will not take a very specific stance on *de re* attitudes and attitude reports.

If we want to give a more precise characterization of the proposed treatment of counterparts, it is helpful to first define a relation which is less restrictive than the strict accessibility relation given in (37). This relation, which I dub 'weak accessibility', holds whenever strict accessibility holds, but in addition it holds across attitude contexts as well:

(47)    *Employment*

Let $\tau(\alpha) = \{\alpha\}$, if $\alpha \in RM_i \cup RM_p$, and $\tau(p+\varphi) = \{p, p+\varphi\}$. Then $Emp(\varphi)$ is the smallest set of terms for which the following hold:

a.    If $P\alpha_1 \ldots \alpha_n \in Con(\varphi)$, then $\cup\{\tau(\alpha_1), \ldots, \tau(\alpha_n)\} \subseteq Emp(\varphi)$

b.    If $\alpha = \beta \in Con(\varphi)$, then $\tau(\alpha) \cup \tau(\beta) \subseteq Emp(\varphi)$

(48)    *Weak accessibility*

$\leqq$ is the smallest preorder for which the following hold:

a.    If $\varphi \leq \psi$, then $\varphi \leqq \psi$

b.    If $p+\psi \in Emp(\varphi)$, then $\varphi \leqq \psi$

For example, $(46b_1) \leqq (46b_2) \leqq (46b_3)$.

The proposed representation of counterparts can be characterized in terms of weak accessibility as follows: if $u \in U(\varphi) \cap U(\psi)$ and $\varphi \leq \psi$, but not $\varphi \leq \psi$, then any pair of values that $u$ can take in $\varphi$ and $\psi$ must be counterparts. Thus, in the present system, there are two ways in which the interpretation of an embedded DRS may be affected by the interpretation of a higher DRS. For any pair of DRSs $\varphi$ and $\psi$: (i) if $\varphi$ is strictly accessible to $\psi$, then the values of any reference markers that $\psi$ shares with $\varphi$ are determined in $\varphi$; and (ii) if $\varphi$ is only weakly (not strictly) accessible to $\psi$, then the values of any reference markers that $\psi$ shares with $\varphi$ must be counterparts.

To ensure that DRSs like (45b) and (46b) get the intended interpretations, we must impose restrictions on the class of eligible embedding functions. The following might be a possible embedding function associated with (45b):

$$\left[ \begin{array}{ccl} x & \to & a \\ y & \to & b \\ p & \to & ... \\ q & \to & \left\{ \begin{array}{l} \langle w_1, [y \to b_1] \rangle \\ \langle w_2, [y \to b_2] \rangle \end{array} \right\} \end{array} \right]$$

This will no longer be an admissible embedding function unless $b_0$ and $b_1$ and $b_0$ and $b_2$ are counterparts. More generally, once a variable has been assigned a value, downstairs occurrences of the same variable should only be assigned values that are counterparts of values assigned upstairs. Or, to put it somewhat less sloppily:

(49)     *Dependents*

         For any embedding function $f$, dps$(f)$ is the smallest set for which the following hold:

         a.     If $p \in$ dom$(f) \cap$ RM$_p$ and $\langle w, g \rangle \in f(p)$, then $g \in$ dps$(f)$

         b.     If $g \in$ dps$(f)$ then dps$(g) \subseteq$ dps$(f)$

(50)     *Admissible embedding functions*

         Let $\approx$ be the counterpart relation fixed by a given model M, and $f \in F_M^n$, for some $n \geq 0$. Then an embedding function $f$ is *admissible* iff the following hold:

         a.     If $g \in$ dps$(f)$ and $\alpha \in$ dom$(f) \cap$ dom$(g)$, then $f(\alpha) \approx g(\alpha)$

         b.     If $g \in$ dps$(f)$ then $g$ is an admissible embedding function

Finally, we stipulate that the interpretation of a DRS in a model M may only involve embedding functions that are admissible in M.

## 5.5 Presuppositions in and of belief reports

Although I have shown how belief reports can be represented in DRT, I have said nothing yet about the mechanisms that produce these representations. These mechanisms are the subject of the present section and the next one. I have proposed that the semantic representation of a sentence of the form '*a* believes that S' involves the following ingredients:

- a reference marker *u* which represents *a*;
- two propositional reference markers *p* and *q*;
- conditions of the form '*u* believes *p*' and '*u* believes *q*' which express that *p* and *q* represent *a*'s doxastic context;
- and, finally, a condition of the form '$q = p+\varphi$', where $\varphi$ reflects the semantic content of S.

I now add to this scheme one decisive qualification, namely that the content of a belief report divides into a presupposed and a non-presupposed part: '*a* believes that S' triggers the presupposition '*u* believes *p*', and asserts '$q = p+\varphi$' and '*u* believes *q*'. In words: whenever a speaker ascribes to an individual *a* the belief that S, he presupposes that *a* has beliefs and asserts that those beliefs entail S.

That '*a* believes that S' presupposes that *a* has beliefs is confirmed by the standard diagnostic tests. For example, if a speaker utters a sentence of the form '*a* does not believe that S', he obviously takes it for granted that *a* believes something, and this implication can be suspended in a way that is characteristic of presuppositions. Compare the following examples for instance:

(51)    a.    It is possible that my Powerbook believes that the square root of 4 is 2.

          b.    If my Powerbook has any beliefs at all, it still doesn't believe that the square root of 4 is 2.

Normally, we would be entitled to infer from an utterance of (51a) that the speaker believes that his Powerbook has beliefs, and the same inference is blocked in (51b). Of course, this is precisely what we would expect if this inference were of a presuppositional nature. Furthermore, the assumption that '*a* believes that S' presupposes that *a* has beliefs is made, albeit tacitly, by Heim, too. True, it doesn't do any work in her theory, but it is encoded in her semantics of *believe*. Within the present framework, this assumption will prove to be crucial, and this is because the presupposition induced by '*a* believes $\varphi$' consists of a propositional term representing *a*'s doxastic context. I have shown that such terms may affect the accessibility structure of a DRS, and indeed we will see in a moment that the *projection* characteristics of the

verb *believe* are determined, inter alia, by the presupposition it *induces*; in §
5.7 the same will be seen to hold for *want*, too.

   If a speaker employs an expression or construction that triggers a
presupposition $\chi$, he indicates that he takes $\chi$ to be given in the context in
which it occurs — i.e. $\chi$ must be given in one of the DRSs that are accessible
to $\chi$'s home DRS. But meanwhile the notion of context has become an
ambiguous one: there is a narrow and a wide notion of context, defined by
strict and weak accessibility, respectively. However, there are *a priori*
grounds to expect that it is the wider notion of context which is relevant in
this connection. In an important sense, *everything* that is contained in a DRS
$\varphi$ represents information that the speaker is responsible for. This holds for
belief contexts, too: if $\varphi \leq \psi$ but not $\varphi \leq \psi$, and $\psi$ represents somebody's
belief, then it still is a representation that the speaker is committed to. It
is the speaker who selects the linguistic means for characterizing a
subject's beliefs, and in particular it is the speaker who chooses between
presuppositional and non-presuppositional devices.[11] It would have been
remarkable, therefore, if presuppositions were confined to the narrow
context in which they are triggered, and it is only natural to assume that
presupposition projection is constrained only by weak accessibility. At any
rate, I will assume it to be so, and unless explicitly indicated otherwise, I shall
in the following use 'accessibility' as short for 'weak accessibility'.

   Assuming that the lexical entry of *believe* is in fact structured as I claim it
to be, and taking the binding theory as presented in Chapter 2, we get a
number of interesting predictions about presuppositions in attitude contexts.
To begin with, we predict that presuppositions may be bound in *believe-
believe* sequences. The following is a case in point:[12]

(52)      a.      If Barney$_i$ believes that he$_i$ has a son$_j$, he$_i$ also believes that he$_j$
                    is a gifted ventriloquist.

          b.      [: [p, p': B believes p, B believes p', p' = p+[x: x is B's son]]
                    $\Rightarrow$ [q, q': B believes q, B believes q',
                              q' = q+[z: ventriloquist z]]]

---
[11] As it turns out, this point is easily misunderstood, perhaps because the term 'presupposition'
is standardly used with at least two different meanings, which are sometimes run together. On
the one hand, it is customary to say that a given expression triggers or induces a presupposition,
while, on the other hand, the same term is used to characterize certain commitments that the
speaker incurs as a consequence of his saying something. So even if a speaker is responsible for
choosing an expression $\alpha$ that triggers the presupposition (in the first sense) that $\varphi$ is true, it does
not follow that he cannot use $\alpha$ without presupposing (in the second sense) that $\varphi$ is true. (If this
followed, presupposition projection would not be problem.)
[12] Here and in the following I use capital letters as semantic proxies for names. This is merely
to avoid a proliferation of reference markers.

I have said that '*a* believes that S' triggers the presupposition that *a* has beliefs. Two such presuppositions are triggered in (52b): one in the antecedent of the conditional, and one in the consequent. The binding theory predicts that the first of this pair will be accommodated in the main DRS, because there is no suitable antecedent available, and there is nothing to block global accommodation. This gives us (52c):

(52)  c.   [p: B believes p,
             [p': B believes p', p' = p+[x: x is B's son]]
             ⇒ [q, q': <u>B believes q</u>, B believes q',
                          q' = q+[z̲: ventriloquist z]]]

The second presupposition is bound in the antecedent of the conditional, as a result of which we get (52d):

(52)  d.   [p: B believes p,
             [p', q: q = p', B believes p', B believes q,
              p' = p+[x: x is B's son]]
             ⇒ [q': B believes q', q' = q+[z̲: ventriloquist z]]]

Once q has been bound to p', the first belief DRS is accessible to the second, and therefore the reference marker x is now accessible to the anaphor z, and the latter may be bound to the former. This gives us (52e), which is equivalent to (52f):

(52)  e.   [p: B believes p,
             [p', q: q = p', B believes p', B believes q,
              p' = p+[x, z: z = x, x is B's son]]
             ⇒ [q': B believes q', q' = q+[: ventriloquist z]]]

        f.   [p: B believes p,
             [p': B believes p', p' = p+[x: x is B's son]]
             ⇒ [q': B believes q', q' = p'+[: ventriloquist x]]]

Which, as I have argued in the preceding section, is an adequate rendering of a possible reading of (52a).

This example already reveals some of the explanatory potential of the theory. First, the theory predicts that (52a) presupposes that Barney has beliefs. Secondly, the theory explains how the two separate belief reports in (52a) come to refer to the same belief context. Not that this is a particularly hard problem to solve, but it should be noted that, in the present framework, its solution requires no special provisions or assumptions. Thirdly, and most importantly, the theory offers a straightforward explanation of the fact that an anaphor in one belief report can be bound to an antecedent in another. This explanation, naturally, extends to presuppositional expressions in

general: if a presupposition can be bound in the belief context in which it arises, the binding theory predicts that this will in fact happen.

If a presupposition $\chi$ is triggered in a belief report on a subject $a$, and $\chi$ cannot be bound, we predict that $\chi$ will be accommodated in the main DRS. For example, it is predicted that, by default, (53a) will be interpreted as presupposing that there is a manager:

(53)    a.    Wilbur believes that the manager is a crook.

        b.    [p, q: W believes p & q, q = p+[x: <u>manager x</u>, crook x]]

        c.    [x, p, q: manager x, W believes p & q, q = p+[x: crook x]]

In (53b) the presupposition that there is a manager cannot be bound, and therefore the binding theory predicts that, in the absence of a suitable antecedent, the presupposition is preferably accommodated in the main DRS.

(Actually, I was cheating just now, though not much. If I had blindly applied the rules of the binding theory as defined in Chapter 2, the reference marker x in Wilbur's belief DRS would have become unbound. This problem is liable to raise its head whenever a presupposition is projected to a DRS that is not strictly acessible to its home DRS. But I didn't apply the rules blindly, and inserted a copy of x into the embedded DRS, and will continue to do so whenever the need arises. The binding theory is easily revised so as to sanction this sleight of hand, for instance by adjusting the binding mechanism or the representation of presuppositions.)

Of course, if a presupposition can be accommodated outside of the belief context in which it is triggered, it may also be bound there. The following is a case in point:

(54)    If Barney was tripped, then he believes that it was Fred who tripped him.

Intuitively, the presupposition triggered by the *it*-cleft in the belief report about Barney (that Barney was tripped) is bound in the antecedent of the conditional, and this is precisely what our theory predicts:

(55)    a.    [: [x: x tripped B] $\Rightarrow$
                   [<u>p</u>, q: <u>B believes p</u>, B believes q,
                         q = p+[<u>z</u>: <u>z tripped B</u>, z = F]]]

        b.    [p: B believes p,
                   [x, z: z = x, x tripped B, z tripped B] $\Rightarrow$
                   [q: B believes q, q = p+[z: z = F]]]

        c.    [p: B believes p,
                   [x: x tripped B] $\Rightarrow$ [q: B believes q, q = p+[x: x = F]]]

The first presupposition in (55a) is accommodated in the main DRS because it cannot be bound. The second presupposition, that someone tripped Barney, can now be bound in the antecedent of the conditional. The resulting DRS is (55b), which is equivalent to (55c), and which says that if Barney was tripped, then Barney believes of his counterpart to the hypothetical individual who tripped Barney that he is Fred.

To sum up: the theory that I propose solves a number of problems concerning the behaviour of presuppositions in belief contexts which, taken at face value, might seem to point in opposite directions:

- First, it accounts for the fact that presuppositions can be bound in *believe-believe* sequences, as in (52a).

- Secondly, it predicts that a presupposition may be bound beyond the confines of a belief context, and thus accounts for cases like (54).

- Thirdly, it predicts that, whenever a presupposition cannot be bound, it will, ceteris paribus, be accommodated in the main DRS. For example, (53a) is predicted to presuppose that there is a manager.

## 5.6 Importation

My theory gives qualified support to the e-principle, which implies, speaking in terms of the binding theory, that a presupposition triggered in a belief context will project to the main DRS. The qualification made here is that this will not happen if a presupposition is bound or accommodated at an intermediate level of representation, which may be the belief context in which it was triggered. We have already seen how presuppositions can be bound within belief contexts. The following example illustrates how, under special circumstances, they may be accommodated locally. Suppose that it is part of the common ground that the university of Osnabrück doesn't employ a hangman (which, to the best of my knowledge, is not the case, but that is as it may be). Then one of the interlocutors might nonetheless volunteer:

(56)     Professor Müller believes that the university hangman is after him.

In the given context, this would be construed as 'Professor Müller believes that the university has a hangman, who is after him', which is to say that the general preference for global accommodation is overridden by contextual information, and the presupposition is accommodated locally, within professor Müller's doxastic context.

If a presupposition is neither bound nor accommodated within the belief context in which it is triggered, it will be bound or accommodated externally. However, I have argued that presuppositions triggered in attitude contexts often, though not always, receive a two-sided reading, which has an internal

as well as an external component (§ 5.1). This observation can now be expressed as follows: if a presupposition is neither bound nor accommodated within the belief context in which it is triggered (hence is projected to an external position), it will often be construed internally, to boot. This extra internal construal is not of a presuppositional nature, though it is based upon a presupposition, construed externally. It will be convenient to have a name for this additional inference: I will call it 'importation', the idea being that a piece of information is imported into a belief context. In the following I will elaborate upon this proposal, and contrast it with the opposite view advocated by Karttunen and Heim.

Consider the following example, which I borrow from Heim (1992: 206):

(57)     a.    John believes that it stopped raining.

         b.    John believes that it was raining.

         c.    It was raining.

Some theories of presupposition predict that (57a) presupposes that (57b) is true; this holds, in particular, for the theories proposed by Karttunen (1974) and Heim (1992). For these theories the problem is to explain why, in addition, (57c) would occasionally be conveyed by an utterance of (57a), as well, and here an *ex*portation argument naturally suggests itself. This is in fact the course taken by Karttunen and by Heim. The latter argues as follows:[13]

> [...] assumptions to be accommodated are supposed to be uncontroversial and unsurprising. [...] So when we hear [(57a)] out of the blue, we know two things: first, as a matter of the semantics of this sentence, we know that it requires the presupposition that John believes that it was raining. Second, we know that the speaker takes this to be uncontroversial and unsurprising. Now why would it be unsurprising that John has such a belief? The most natural guess is that it would be unsurprising because it was in fact raining and John was in an appropriate position to find out. Of course, these are not the only possible conditions under which someone might form a belief that it was raining; but they are the most *normal* conditions. (Heim 1992: 212)

---

[13] Kay (1992) presents an argument which is similar to Heim's. This is one of two strategies considered by Heim for explaining two-sided readings (the other, which is based upon the idea that external readings might in fact be *de re* construals, was discussed in § 5.2). The argument that Heim presents in the passage cited here is reminiscent of Karttunen and Peters's proposal for dealing with the proviso problem, which was discussed in Chapter 3.

This argument may seem plausible enough at first, but it doesn't withstand closer scrutiny. First, consider the following example:

(58)    a.    Wilma is polishing her stethoscopes.
          b.    Wilma has several stethoscopes.
          c.    Wilma is a doctor.

Heim's argument would imply that we may reason as follows: (58a) requires the presupposition that (58b) is true, and the speaker takes this presupposition to be unsurprising. How could that be? The most natural guess is that (58c) is true, for normally speaking it is doctors who own stethoscopes. This result is plausible, too. What I find objectionable, however, is that the same argument that has taken us from (57a) to (57c) should take us from (58a) to (58c). For it would seem to follow from this that (58c) stands to (58a) as (57c) stands to (57a), and this is arguably incorrect. (58c) is clearly a bridging inference that is based upon plausibility, just as Heim's argument suggests, and it is easily defeasible: if (58a) is followed by, say,

(59)      She has been collecting them since she was 12 years old.

the inference to (58c) is simply suspended, although (58c) and (59) are not incompatible in any way. But the inference from (57a) to (57c) is not undone so easily: it clearly requires special means to accomplish this, and therefore this is presumably not a defeasible inference to begin with, and at any rate it is not of the same nature as the inference from (58a) to (58c).

    This conclusion is confirmed by the following observation. If we take (57b), which according to the Karttunen-Heim theory is presupposed by (57a), and give it as a complement to a factive verb like *know*, the inference to (57c) evaporates without a trace:

(60)      Fred knows that John believes that it was raining.

This already suffices to prove that there is something seriously amiss with Heim's argument. It is universally agreed that (60) presupposes that John believes that it was raining, and Heim's theory predicts that (57a) has the same presupposition. But then why should a hearer infer that it was raining in one case but not in the other? If Heim's reasoning were sound we should expect it to apply in both cases, but as it clearly doesn't apply in the case of (60), it is dubious that it should hold in the case of (57a). Furthermore, if we embed the presupposition of (58a) in the same way,

(61)      Fred knows that Wilma has several stethoscopes.

the resulting sentence still suggests that Wilma is a doctor. Which confirms our suspicion that (57c) does not stand to (57a) as (58c) stands to (58a). The latter is clearly a bridging inference, whereas, contrary to what Heim's argument suggests, the former is not.

There is a further, more general, objection against Heim's argument: it takes the wrong direction, in that it attempts to draw conclusions about the speaker's beliefs on the basis of the beliefs that he ascribes to a third party, instead of the other way round. Although the notion is notoriously hard to make precise, it is generally accepted that in construing the beliefs of others we operate on a principle of charity: we try to avoid the conclusion that other people's beliefs are contradictory, we credit them with knowledge that we take to be uncontroversial or commonly available, and so on. It is only natural to assume that the same principle underlies our speaking of other people's beliefs. Putting the point without any of the necessary nuances: if a person *a*'s doxastic context is the subject of a conversation, the interlocutors will tend to assume that *a* believes what they believe. This is a caricature but the basic idea is surely right. Heim's argument, however, proceeds in the opposite direction, which is not nearly as plausible: in general, if I describe someone as believing that φ, I intend to at least remain uncommitted concerning the truth of φ. What is more, if someone describes a person as believing (rather than knowing) that φ, his description may often be taken to imply that he doubts that φ is true or is even convinced that it is false. In short: if a speaker reports on the beliefs of a person *a* there may be a tendency to ascribe to *a* beliefs that the speaker is taking for granted, but it is *a priori* unlikely that the hearer will ascribe a certain belief to the speaker because the speaker ascribes it to *a*.[14] It is therefore a mistake to claim, as Kay (1992) does explicitly, that:

> [...] in reporting the thoughts of another, it appears to be the case that unless the speaker states or implies (or it is given by the background) that the speaker thinks the reported thoughts are wrong, the default assumption is that the speaker thinks they are right. (Kay 1992: 365)

---

[14] I have the impression that Heim doesn't distinguish between the two directions of 'information flow' in talk about beliefs, or between exportation and importation, as I have dubbed them. That is at least what her citing the following passage from Karttunen (1973b) suggests:

> [...] unless it has been indicated otherwise, [John] can be assumed to share the speaker's beliefs. In other words, there is a natural spill-over from [the global context to John's belief context]. Consequently, in situations where nothing has been said about [John's] beliefs, one tends to think that, if the presuppositions of [(57a)] are satisfied, they are satisfied by virtue of the speaker's tacit assumption that [John] shares his beliefs. (Karttunen 1973b, as cited by Heim 1992: 184)

(Not having access to the Karttunen manuscript, I quote from Heim substituting Karttunen's example with the one used earlier on.) Note that this passage starts out with an unequivocal appeal to some importation principle, but that the 'spill-over' which the Karttunen-Heim theory requires goes in the opposite direction, from John's belief context to the global context, and is therefore a form of exportation.

It may sometimes happen that, when *a* reports on what *b*'s statements, suspicions, or opinions are, *a* implies that *b* is right. But it is certainly not the normal case.

The upshot of these observations is that the prospects of the importation view are decidedly brighter than those of the exportation view. This is a welcome result for my theory of presupposition, because it leaves a gap which seems to call for some version of importation. I now want to argue that, apart from the fact that it is to be preferred to its opposite, importation is plausible in itself and motivated by independent considerations.

That importation is a plausible strategy is conceded even by Heim, who appeals to it after having suggested that some (in my terminology) external readings of presuppositional expressions are in fact *de re* construals. I have shown that this idea fails in the general case, but it receives seeming support from some presupposition triggers, such as definite NPs. Suppose now that the definite subject NP in (62) is construed *de re*:

(62)    Ralph thinks that the man he saw at the beach is a spy.

Let us say that this sentence is to be read as claiming that there is some acquaintance relation R such that Ralph bears R to the man he saw at the beach, i.e. Bernard J. Ortcutt, and thinks that whoever he bears R to is a spy. If we adopt some account along these lines, it is plausible to surmise that the mere fact that a speaker who utters (62) is referring to Ortcutt with the help of this particular definite description 'will bias the hearer towards the assumption that the intended acquaintance relation between Ralph and Ortcutt is the one established in the beach-encounter.' (Heim 1992: 210) Thus even if (62) is construed *de re*, it is likely to be understood as implying that Ralph believes that he has seen a man at the beach who is a spy. According to this scenario, then, (62) will in a sense be read *de re* as well as *de dicto*.

Even if we must reject the theory in which it is embedded (as I have argued we must), the main thrust of this importation story is in the right direction. Heim's analysis of (62) is a plausible one, although I prefer to view it in terms of the external/internal distinction, and its plausibility even increases when we consider other presupposition triggers than the definite article. Consider, for example, (63a):

(63)    a.    Fred believes that it was Barney who tripped him.

        b.    [p, q: F believes p & q, q = p+[x: x tripped F, x = B]]

        c.    [x, p, q: x tripped F, F believes p & q, q = p+[x: x = B]]

The only presupposition that remains to be processed in (63b) is that someone tripped Fred, and the binding theory predicts that this presupposition will be accommodated in the main DRS, which yields (63c).

This DRS says that an individual *a* tripped Fred, and that Fred believes that some *b*, who is *a*'s counterpart, is Barney. Obviously, if this were all the speaker wanted to communicate, his uttering (63a) would be a pointless thing to do, and it is natural to infer that, according to the speaker, Fred actually believes that *b* tripped him.

No matter what the details of this explanation will look like, we will need something along these lines for independent reasons. For example:

(64)     Fred was tripped by somebody, and he believes that it was Barney.

It is clear that, normally speaking, someone who utters (64) takes it for granted that Fred believes that he was tripped. However, the reading we (and not only we) actually predict for this sentence is (63c), and since this reading doesn't imply that Fred believes that he was tripped, importation is required. Similarly:

(65)     It is raining but Fred believes that it will soon stop.

Fred believes that *what* will soon stop? The answer is obvious, but it requires that the information that it is raining is imported into Fred's doxastic context.

Hence, anaphoric pronouns that occur in attitude contexts provide independent support for the notion of importation. Further independent evidence is furnished by specific indefinites:

(66)     Wilma believes that a boy scout has fallen in love with her.

The indefinite noun phrase in (66) can be construed *de re*, although this is probably not the preferred option.[15] But even if we understand an utterance of this sentence as implying that there is a boy scout such that Wilma believes that he is in love with her, it is rather likely that we would infer that Wilma, too, believes that the individual in question is a boy scout. I have the impression that such two-sided readings are the rule rather than the exception — i.e. *if* an indefinite noun phrase is construed *de re*, it will in general be construed *de dicto* as well. Be that as it may, for my present purposes it is already sufficient that (66) *may* be read as implying both that

---

[15] Some people have played with the idea of analysing specificity in presuppositional terms, the suggestion being that, if an indefinite receives a specific interpretation, it triggers a presupposition that is not given and must therefore be accommodated. This is not a good idea however. Given that presuppositions are defined as pieces of information that are (pretended to be) given, and that accommodation is therefore a repair strategy, a presupposition that *eo ipso* is not given and must be accommodated is simply a contradiction in terms. So it is unlikely that specific indefinites can be accounted for by a theory of presupposition. Note that this point does not presuppose that the binding theory is correct. It assumes merely that presuppositions are (presented as) given information, which is an uncontroversial premise. (It is accepted by satisfaction theorists, too. The main issue between the binding theory and the satisfaction theory is over the question what givenness means in this connection.)

there is a boy scout such that Wilma believes that he is in love with her, and that Wilma, too, believes that the individual in question is a boy scout. For if (66) may be read this way, it obviously isn't sufficient if we can explain how the indefinite in (66) can be construed *de re*. What we need, in addition, is an account of the fact that it can be read *de dicto* at the same time, i.e. we require some form of importation.

It is evident that importation isn't anything like a logical rule of inference: it doesn't always apply. What is more, I see no good reason to assume that it is default rule, either.

(67)  a.  My mother thought ⎫
      b.  The policeman thought ⎬ that my sister was drunk.

We would normally infer from (67a) that the speaker's mother believed of the speaker's sister that she is his sister. (67b), in contrast, would not normally license the inference that the policeman had this belief. If importation were a default rule, the former would be the 'normal' case, and the latter would be an exception. But this is counterintuitive. I conclude, therefore, that importation is neither a logical nor a default rule: it is just a convenient label for a certain class of context-dependent plausibility inferences.

If this conclusion is correct, then it is further evidence against Heim's theory of belief reports. On Heim's analysis, a sentence of the form '*a* believes that $\varphi\{\chi\}$' will *normally* give rise to the (allegedly presuppositional) inference that '*a* believes that $\chi$'; special mechanisms are called for if we want to explain the (allegedly non-presuppositional) inference that, according to the speaker, $\chi$ is true. But the argument of the previous paragraph strongly suggests that it is precisely the other way round, as indeed my analysis has it.

## 5.7 'Want'

Having thus far confined my attention to the semantics and presuppositional behaviour of *believe*, I want to show in this section how the theory can be extended so as to account for *want*, too. As in the case of *believe*, the semantics of *want* that I adopt is the same as Heim's, and I shall merely complement it with an assumption concerning the presuppositions that this verb triggers. And as with *believe*, the projection properties of *want* will be argued to follow from this.

As we have seen in § 5.2, the following is (a simplified version of) Heim's proposal for the context-change semantics of *want*, which was inspired by Stalnaker's analysis:

(68)  $c[a \text{ wants } \varphi] = \{w \in c \mid \text{dox}_a(w)[\varphi] <_{a, w} \text{dox}_a(w)[\neg\varphi]\}$  (= (24))

Where $<_{a,\,w}$ represents an ordering which ranks propositions according to their desirability for an individual $a$ in a world $w$. What this says is that $a$ wants $\varphi$ iff he prefers all doxastic alternatives in which $\varphi$ is true to any of his doxastic alternatives in which $\neg\varphi$ is true.

This analysis can be transposed into our DRT framework as follows. To begin with, we note that Heim's rule for *want* involves three sets of worlds: a set of doxastic alternatives $dox_a(w)$, the subset of $dox_a(w)$ in which $\varphi$ is true, and the subset of $dox_a(w)$ in which $\neg\varphi$ is true. The third set is obviously derivable once the first two are given. Accordingly, the DRT representation that I suggest involves two propositional reference markers, one corresponding to Heim's '$dox_a(w)$', and one corresponding to her '$dox_a(w)[\varphi]$'. These two reference markers, $p$ and $q$, are related through a condition of the form '$q = p+\varphi$'. The propositional reference marker $p$ corresponds with Heim's '$dox_a(w)$'; in order to bring this out, a condition of the form '$u$ considers $p$' is imposed on $p$, which is to say that the information in $p$ is compatible with $u$'s beliefs (see (42b) for the precise definition of *considers*). Finally, we need to represent the information that, among the $p$-worlds, $u$ prefers the ones that are in $q = p+\varphi$ to any world that is not. To this end, we introduce a three-place relational constant *prefers*, which is interpreted as follows:

(69)    $I_w(\text{prefers}) = \{\langle a, \sigma, \sigma'\rangle \mid w(\sigma') <_{a,\,w} (w(\sigma) - w(\sigma'))\}$

This is a relation which holds between an individual $a$ and a pair of indexed propositions $\sigma, \sigma'$, iff $a$ prefers all worlds in $w(\sigma')$ to any world which is in $w(\sigma)$ but not in $w(\sigma')$. This is more general than what we actually need, but in our encoding of *want*, $w(\sigma)$ will be a set of doxastic alternatives and $w(\sigma')$ that subset of $w(\sigma)$ in which some proposition holds, so this gives us what we want. Instead of 'prefers $u\ p\ q$', I will usually write '$u$ prefers$_p\ q$', which may be read as: 'among the $p$-worlds, $u$ prefers the $q$-worlds to the non-$q$-worlds'.

To summarize: a sentence of the form '$a$ wants S' will be represented by a DRS of the form $[u, p, q: u$ considers $p, q = p+\varphi, u$ prefers$_p q]$, where $u$ is a reference marker representing $a$, $p$ and $q$ are propositional reference markers that represent sets of alternatives that are compatible with $a$'s doxastic context, and $\varphi$ is a DRS that represents the content of S. (70) gives an example:

(70)    a.    Professor Müller wants to play the accordion.

        b.    [x, p, q: PM x, x considers p,
                q = p+[x: x plays the accordion], x prefers$_p$ q]

The underlining in (70b) already reveals what is the keystone of my proposal: analogously to *believe*, the content of *want* is partitioned into a presupposed and a non-presupposed part. (70a) takes some set of alternatives as given

and asserts of this set, represented here by the reference marker p, that professor Müller prefers all p-worlds in which he plays the accordion to any p-world in which he does not. The arguments in favour of this analysis parallel the ones I adduced in support of my analysis of *believe*. In particular, it is easily shown that it is corroborated by the standard diagnostics for presuppositions.

Let us now see what predictions this hypothesis yields.

(71)    a.    Kurt believes that he is being shadowed by a ghost. He wants to write a book about it (the ghost).

        b.    $[p, p', q, q':$
           K believes $p$ & $p'$, $p' = p+$[x: ghost x, x shadows K],
           <u>K considers q</u>, K prefers$_q$ $q'$,
           $q' = q+$[y, <u>z</u>: book y, K writes y about z]]

We assume that the first sentence of (71a) has been processed, and start out from the representation in (71b). This DRS contains presuppositions induced by the verb *want* (q) and the pronoun *it* (z); all other presuppositions are either ignored or assumed to have been processed already. Since *believes* entails *considers*, q may be bound to $p'$.[16] This yields (71c), which is equivalent to (71d):

(71)    c.    $[p, p', q, q': q = p',$
           K believes $p$ & $p'$, $p' = p+$[x: ghost x, x shadows K],
           K considers q, K prefers$_q$ $q'$,
           $q' = q+$[y, <u>z</u>: book y, K writes y about z]]

        d.    $[p, p', q':$
           K believes $p$ & $p'$, $p' = p+$[x: ghost x, x shadows K],
           K prefers$_{p'}$ $q'$, $q' = p'+$[y, <u>z</u>: book y, K writes y about z]]

Now the remaining anaphor, i.e. z, has gained access to Kurt's belief DRS, and is bound to x. The outcome of this operation is (71e), which is equivalent to (71f):

(71)    e.    $[p, p', q':$
           K believes $p$ & $p'$, $p' = p+$[x, z: z = x, ghost x, x shadows K],
           K prefers$_{p'}$ $q'$, $q' = p'+$[y: book y, K writes y about z]]

        f.    $[p, p', q':$
           K believes $p$ & $p'$, $p' = p+$[x: ghost x, x shadows K],
           K prefers$_{p'}$ $q'$, $q' = p'+$[y: book y, K writes y about x]]

---

[16] Strictly speaking, '*a* believes σ' entails '*a* considers σ', in a given world *w*, only if $dox_a(w) \neq \emptyset$. I assume here that this condition is satisfied.

This analysis shows how an anaphoric pronoun in a *want* context can pick up its antecedent from a *believe* context. Similarly, in (72a) the presupposition that Fred's wife is given to drink, which is triggered by the aspectual verb *stop*, is bound in the belief context in the antecedent.

(72)    a.    If Fred believes that his wife has taken to drink again, then he will want her to stop drinking.

       b.    If Fred's wife has taken to drink again, then he will want her to stop drinking.

In (72b), on the other hand, the same presupposition is bound in the antecedent of the conditional, and this is predicted, too, since a *want* context is the same type of entity as a *believe* context, and we have seen already that the theory allows presuppositions to escape from such contexts. By the same token, it is predicted that, if a presupposition triggered in a *want* context cannot be bound, it will be accommodated, which means, as a rule, that it is accommodated globally. This is correct, too, as witness examples like the following:

(73)    Fred wants his wife to stop drinking.

Under normal circumstances, this will be taken to imply that, according to the speaker, Fred's wife has been drinking, which is what I predict.

    We have seen that Heim's theory cannot account for presupposition filtering in *want-want* sequences (§ 5.2). Mine can. To illustrate this, suppose that the discourse in (71a) is continued with an utterance of (74a). Our initial DRS will then be as in (74b), which consists of (71f) incremented with the information furnished by the new sentence:

(74)    a.    And he (Kurt) wants to dedicate it (the book) to Professor Müller.

       b.    $[p, p', q', \underline{r}, r':$
          K believes $p$ & $p'$, $p' = p+[x:$ ghost $x$, $x$ shadows K$]$,
          K prefers$_{p'}$ $q'$, $q' = p'+[y:$ book $y$, K writes $y$ about $x]$,
          <u>K considers $r$</u>, K prefers$_r$ $r'$, $r' = r+[\underline{z}:$ K dedicates $z$ to PM$]]$

In (74a) the neuter pronoun in the scope of *want* is anaphoric on *a book* in (71a). (74a) presupposes that there is a set of indexed worlds which is compatible with the doxastic alternatives that are open to Kurt, and the DRS in (74b) offers three possible antecedents for this presupposition. However, there will be no suitable antecedent for *it* unless $r$ picks up $q'$, which makes available $y$ as an antecedent for $z$. The binding theory predicts that the latter alternative will be preferred, which yields the representation in (74c), or, equivalently but slightly more succinctly, (74d):

(74)  c.  [p, p′, q′, r, r′: r = q′,
K believes p & p′, p′ = p+[x: ghost x, x shadows K],
K prefers$_{p′}$ q′,
q′ = p′+[y, z: z = y, book y, K writes y about x],
K considers r, K prefers$_r$ r′,
r′ = r+[: K dedicates z to PM]]

d.  [p, p′, q′, r′:
K believes p & p′, p′ = p+[x: ghost x, x shadows K],
K prefers$_{p′}$ q′, q′ = p′+[y: book y, K writes y about x],
K prefers$_{q′}$ r′, r′ = q′+[: K dedicates y to PM]]

The results of our analysis of *want* may be summed up in four points:

- First, it predicts that presuppositions may be bound in *believe-want* sequences, as in (71a).
- Secondly, it accounts for the fact that presuppositions may be bound in *want-want* sequences, as when (71a) is followed by (74a).
- Thirdly, a presupposition may be bound beyond the confines of a *want* context, as in (72b).
- Fourthly, if a presupposition triggered in a *want* context cannot be bound it will in general be accommodated at top level, as in (73).

The binding theory doesn't account for all inferences that may be associated with presuppositional expressions in *want* contexts.

(75)  a.  Ralph wants to inform against the man he saw at the beach.

b.  Ralph wants to inform against an individual whom he believes to be a man that he saw at the beach.

Typically, an utterance of (75a) would be taken to imply that the speaker holds (75b) to be true, as well. I agree with this observation, but I am bound to deny that this inference is a presupposition. According to the analysis that I have presented, (75a) presupposes (in the absence of contextual information to the contrary) that Ralph saw a man at the beach. This is intuitively correct, or at least it accounts for an inference that would normally be licensed by an utterance of this sentence. The additional inference yielding (75b) is not something a theory of *presupposition* can explain, but in the last section we have seen how such inferences can be secured in the case of *believe*. What I said there applies here, too: there is an independently motivated plausibility inference that, on the basis of an external construal of the presupposition triggered by the definite NP in (75a), yields the conclusion that (75b) is true, as well. As in the case of *believe*, this independent evidence is furnished by anaphoric pronouns and specific indefinites occurring in attitude contexts.

(76)     a.    Fred will be tripped by somebody, and he wants it to be
                Barney. (cf. (64))

         b.    It is raining, and Fred wants it to stop. (cf. (65))

(77)           Wilma wants a boy scout to make love to her. (cf. (66))

In (76a), the pronoun *it* occurs in the context created by 'Fred wants . . .', and resolving this pronoun is not in itself sufficient to secure the inference that, according to the speaker, Fred believes that someone tripped him. But clearly this is something we would normally infer from an utterance of (76a).The same holds, mutatis mutandis, for (76b). Similarly, even if the indefinite in (77) is construed *de re* and given wide scope, we would normally infer that Mary believes of the individual with whom she wants to engage in amorous transactions that he is a boy scout. In either case importation appears to be called for, and thus these observations provide independent support for the sort of inference that is needed to round out our presuppositional analysis of examples like (75a).

CHAPTER 6

# Modals

As it stands, the binding theory does not account for (1a), although it does offer an explanation for (1b):

(1)    a.    Maybe Barney owns a Chevrolet and maybe he keeps it hidden somewhere.

        b.    Maybe Barney owns a Chevrolet and maybe he keeps his Chevrolet hidden somewhere.

The presupposition triggered by *his Chevrolet* in (1b) cannot be bound, and will therefore have to be accommodated, and since global accommodation would violate Gricean constraints on interpretation, it will be accommodated locally. This procedure yields a reading that is intuitively correct but it does not work for (1a), because even if the presupposition triggered by the pronoun could be accommodated, there is nothing to prevent global accommodation in this case, and even if the presupposition were accommodated locally, we would still not obtain the correct reading for this sentence, which should come out equivalent to (1b).

If the accommodation account does not work for (1a), there is good reason to suspect that it shouldn't apply to (1b), either. This suspicion is strengthened by the intuition that in (1a) as well as in (1b), the second modal somehow picks up and extends the hypothetical context created by the first, and that the reference marker introduced by the indefinite noun phrase in the first conjunct should be accessible whenever this context is being extended, which is what the second conjunct does. Many of the authors who have discussed this type of example, beginning with Karttunen (1976), appear to share this intuition, which implies that the personal pronoun *it* in (1a) and the definite noun phrase *his Chevrolet* in (1b) are *bound* rather than accommodated.

The phenomenon illustrated by (1a) has come to be known under the name of 'modal subordination', which was coined by Roberts (1987, 1989, 1996). Initially Roberts restricted the notion of modal subordination to anaphoric pronouns, but as she observes in her 1996 paper, the same phenomenon occurs with all sorts of presuppositional expressions, as (1b) illustrates (this type of example was discussed already by Gazdar 1979). We

will see however that Roberts's analysis of modal subordination does not fit very comfortably with the fact that modal subordination occurs with presuppositions in general, and not just with the presuppositions triggered by pronouns. On the other hand, I will argue that Roberts overextends the notion of modal subordination in certain ways, for example because she uses it for analysing disjunctions.

Roberts's theory is discussed and criticized in the first section of this chapter, which also serves to present some of the puzzles concerning modal subordination (§ 6.1). In the subsequent sections I present an alternative account, which is much in the spirit of Roberts's proposal (§§ 6.2 and 6.3). In fact, I almost completely agree with the *informal* account of modal subordination that Roberts offers. The principal ingredient that is lacking from her account, in my opinion, is the idea that modal expressions are presupposition inducers. Once this element is put into place, modal subordination is automatically accounted for by the binding theory. This analysis of modals, which is a straightforward extension of the theory proposed in the previous chapter, diverges considerably from the *letter* of Roberts's proposal, though. In particular, it will turn out that modal subordination can be explained without resorting to the extremely powerful mechanism of 'antecedent accommodation' which is crucial to Roberts's theory, but which is also its weak spot, as I will argue. My account of modal subordination applies to conditional sentences, too, which are the subject of § 6.4. In the course of this chapter it will turn out that 'modal subordination' is actually a misleading term, because the same phenomenon can be observed in attitude reports (as we have seen in the preceding chapter), and with all sorts of quantifier expressions. By way of summing-up, the last section documents this pervasive pattern (§ 6.5).

## 6.1 Modal subordination

Consider the following example:

(2)      A thief might break into the house. He might take the silver.

The most natural reading of (2) is the one on which the modal in the first sentence takes wide scope with respect to *a thief*, so we obtain a reading for the discourse as a whole which may be paraphrased as follows: 'It might happen that a thief broke into the house and that this thief took the silver.' It is the programme of this chapter to account for this intuition and to explain how it can be that the personal pronoun in the second sentence is anaphorically dependent upon the indefinite in the first.

Let me start by briefly considering two straw attempts at executing this programme, just to get them out of the way. First, it might be thought that (2)

can be dealt with by extending the scope of the first modal. The second *might* is then construed as a flag which indicates that the hypothetical context introduced by the first is being taken up again. Thus we start out from a structure as in (3a), which gets reanalysed into (3b), after which the anaphoric reference marker z can be identified with its antecedent, as desired. (3c) represents the final reading of (2):

(3)    a.    [: ◇[x: thief x, break-in x], ◇[z̲: take-silver z]]

       b.    [: ◇[x, z̲: thief x, break-in x, take-silver z]]

       c.    [: ◇[x, z: z = x, thief x, break-in x, take-silver z]]

This proposal has two attractive features: it captures the intuition that the second modal in (2) extends the hypothetical state of affairs brought up by the first, and it allows us to treat the personal pronoun in the second sentence as a regular anaphor, which is what we wanted. But unfortunately, it is fairly obvious that this strategy will work only as long as the modals that we are dealing with are in the same DRS, which need not be the case, as the following example shows:

(4)    If it is possible that a thief breaks into the house, it is also possible for him to take the silver.

Here it is not possible, obviously, to merge the two modal contexts, and therefore our first proposal quickly breaks down.

One might think that perhaps the problem can be solved by copying over the material in the scope of the first modal into that of the second. This procedure allows us to deal with example (2), and it also gives acceptable predictions for (4).

(5)    a.    [: [: ◇[x: thief x, break-in x]] ⇒ [: ◇[z̲: take-silver z]]]

       b.    [: [: ◇[x: thief x, break-in x]] ⇒
             [: ◇[x, z: z = x, thief x, break-in x, take-silver z]]]

In comparison with the first attempt this one strikes me as rather clumsy, precisely because it requires that material be copied from one corner in a DRS to another. But apart from that it doesn't work much better than our first strategy, as the following discourse demonstrates (this example is Roberts's):

(6)    A thief might break into the house. He would take the silver.

Here the two modals are in the same DRS, but this time they don't have the same force, which makes it impossible to merge them. Furthermore, copying the material in the scope of *might* into the scope of *would* doesn't give an acceptable result either, for then we obtain a reading which says, in effect, that in some sense of 'necessary', it is necessary that there will be a thief that

breaks into the house and takes the silver, which is not what the second sentence in (6) means. Rather, this sentence seems to say something like: '*If a thief broke into the house, he would take the silver.*'

It is a familiar observation that the meaning of a modal expression is dependent on contextual factors. This context dependence is quite evident in examples like the following:

(7)       Your teeth might fall out.

In this example it is perfectly clear what kind of contextual information the sentence requires, at least intuitively speaking: (7) means something like 'If the circumstances were to be such and such, your teeth might fall out,' and unless the context fixes what 'such and such' means, the sentence will be unintelligible. Consequently, it is difficult to imagine a conversation opening with an utterance of (7). If we put together this insight with the standard construal of modals as quantifiers over possible worlds, we cannot but conclude that the quantificational domain of a modal expression is restricted by the context in which it occurs.[1]

This context dependence is Roberts's starting point. She observes that the second modal in (6) tends to be read as a conditional whose antecedent makes explicit the domain of the modal *would*. This is in conformity with what we said in the preceding paragraph. What Roberts adds to this picture is a mechanism that actually fills in the modal domain with material furnished by the context. She calls this mechanism 'antecedent accommodation'.[2] To illustrate the workings of this mechanism, let (8) be the initial representation of (6).

(8)       [: $\Diamond$[x: thief x, break-in x], [ ... : ... ] $\Box\rightarrow$ [z: take-silver z]]

In this DRS the first sentence of (6) is represented as a simple possibility statement. This is not to say that this sentence is not relativized to a modal domain, but Roberts argues (1989: 698) that in this particular case the more orthodox representation is adequate enough, and I shall assume for the time being that she is right about this. The second sentence of (6) is represented initially by a conditional whose antecedent has not yet been specified. The connective '$\Box\rightarrow$' is interpreted as a kind of conditional necessity, and its interpretation makes the antecedent accessible to the consequent, as does the ordinary DRT conditional. Now antecedent accommodation as presented by Roberts fills in the antecedent of such a conditional, generally by copying over information from that part of the DRS which represents the immediately preceding discourse. In the example under discussion this

---

[1]   See Lewis (1979) and Kratzer (1981, 1991). The latter traces back the key observations to C.S. Peirce.

[2]   In her 1996 article, Roberts presents this as an instance of local accommodation.

means that the material in the scope of the first modal is copied into the antecedent of the second, which makes the reference marker x accessible to z, so that z can be bound as follows:

(9)     [: $\diamondsuit$[x: thief x, break-in x],
        [x, z: z = x, thief x, break-in x] $\square\!\!\rightarrow$ [: take-silver z]]

Simplifying somewhat, (9) says that there is a possible world in which a thief breaks in, and that in all possible worlds in which a thief breaks in, he takes the silver. Thus antecedent accommodation allows us to assign adequate truth conditions to (6), and furthermore, this can be done in a way that accounts for the intuition that the pronoun in (6) is an ordinary anaphor. Of course, the same account will work for the examples discussed earlier on.

Before I set out to explain what, in my opinion, is wrong with Roberts's proposal, I want to emphasise that I agree with what I take to be her fundamental claim, namely that modal subordination must be explained in terms of the interaction between a modal domain and the context. What I disagree with is her employing the notion of antecedent accommodation to get to grips with this interaction. I shall argue later that a modal expression presupposes its domain, and attempt to show that on this assumption modal subordination is automatically accounted for the binding theory. Antecedent accommodation, in contrast, is a device that is purely ad hoc, and excessively powerful to boot.

Roberts's notion of antecedent accommodation is quite different from the notion of accommodation as it is employed in the binding theory. Our notion of accommodation is a comparatively weak one, because it is restricted in two ways. First, it is always a presuppositional expression that determines what is to be accommodated. For example, if a speaker uses a definite noun phrase of the form 'the N', and the audience don't know of any N, then they may agree to accept that some N exists. Thus it is always the content of the presuppositional expression that is accommodated.[3] Furthermore, we have seen that both the type of presuppositional expression and its content determine whether or not it can be accommodated at all. For example, poor presuppositions, as induced by pronouns and other semantically attenuate definites, cannot be accommodated, generally speaking (§ 2.5).

Roberts (1989: 699) concedes that hers is an 'extension' of the original concept of accommodation. She doesn't say exactly what she means by this, but her use of the notion suggests that in antecedent accommodation, it is the *context* that determines what material is to be accommodated. As Roberts

---

[3]   This is in accordance with Stalnaker's remarks on accommodation. Recall from § 1.2 that the notion of accommodation goes back to Stalnaker (1973, 1974) and Karttunen (1973). Lewis (1979) gave accommodation its name but at the same time widened the Stalnaker/Karttunen concept considerably, too, as we will presently see.

sees it, in modal subordination the modal expression merely indicates that an antecedent expression needs to be specified. It is the context that is made responsible for this specification, and there are no principled restrictions on what can and cannot be accommodated. In this respect, she follows the lead of Lewis (1979) and Kratzer (1981), who also call upon accommodation to explain domain restriction in modals, and who seem to have a similarly unrestricted notion of accommodation in mind.

Antecedent accommodation is, as Roberts (1989: 701) herself observes, 'a very powerful device'; Kratzer (1981: 61) candidly refers to it as 'black magic'. What makes antecedent accommodation such a powerful device is that the context is being held responsible for procuring 'missing' information. It is only to be expected that such a mechanism is going to 'work' also in cases in which we would not expect it to, as in the following:[4]

(10)    Either there's no bathroom in this house or it's in a funny place.

Somewhat surprisingly, Roberts diagnoses this as an instance of modal subordination, too, and proposes the following analysis:

(11)    a.    [: [: ¬[x: bathroom x, in-this-house x]] ∨ [z: in-funny-place z]]

        b.    [: [: ¬[x: bathroom x, in-this-house x]] ∨
              [[x, z: z = x, bathroom x, in-this-house x] □→
                 [: in-funny-place z]]]

Here the second disjunct in (11a) is reanalysed into a conditional whose antecedent is furnished by the first disjunct. I don't think this is correct. I have argued that this type of example should be analysed in terms of bridging (§ 2.4.2), but apart from that I fail to see why it should be treated as an instance of modal subordination, for it is anything but obvious that the interpretation of (10) involves a modal element at all. Roberts's proposal implies that (10) is on a par with:

(12)    Maybe there's a bathroom in this house and maybe it's in a funny place.

This is a genuine instance of modal subordination, and it differs from (10) in at least two ways. First, in the case of (12) there is a clear intuition that the hypothetical context set up in the first conjunct is extended in the second half of the sentence, whilst in the case (10) there is no such intuition, of course. Secondly, and in conformity with the previous point, (10) is quite unlike (12) in that the order between the presupposition and its antecedent doesn't matter: (13a) is about as good as (10), and (13b) surely is fully acceptable, whereas (14a, b) are both equally infelicitous (on the intended reading):

---

[4]    Roberts attributes this particular example to Barbara Partee, but examples of this type were discussed already by Geach (1962). Cf. §§ 2.3, 2.4.2, and 4.4.

(13)   a.   Either it is in a funny place or there is no bathroom in this house.

       b.   Either the bathroom is in a funny place or there is no bathroom in this house.

(14)   a.   Maybe it's in a funny place and maybe there's a bathroom in this house.

       b.   Maybe the bathroom is in a funny place and maybe there's a bathroom in this house.

The main reason why I have brought up this subject is not to show that Roberts's treatment of disjunction misguided. It is rather because the very fact that this treatment is possible at all vividly illustrates that antecedent accommodation is a panacea. To drive home this point, consider the following variation on (10):

(15)   Either there's no bathroom in this house or I just cannot find it.

If (11) is an acceptable analysis of (10), then Roberts's theory can handle (15), too. Let (16a) be the initial representation of (15):

(16)   a.   [: [: ¬[x: bathroom x, in-this-house x]] ∨ [: ¬[z̲: I can find z]]]]

       b.   [: [: ¬[x: bathroom x, in-this-house x]] ∨
            [[x, z: z = x, bathroom x, in-this-house x] □→
               [: ¬[: I can find z]]]]]

       c.   [: [: ¬[x: bathroom x, in-this-house x]] ∨
            [: ¬[: [x, z: z = x, bathroom x, in-this-house x] □→
               [: I can find z]]]]]

In order to obtain (16b), we have applied the same procedure as in (11): the second disjunct is turned into a conditional whose antecedent is the negation of the first. But in this case we seem to get at least one further interpretation into the bargain. The disjunction in (15) differs from (10) in that its second member *contains* a structure that can also be subjected to modal subordination, and there is nothing to block the reading in (16c), where antecedent accommodation has applied to the negated material in the second disjunct of (16a). Since (16b) and (16c) are not equivalent, it appears as if Roberts's theory predicts that (15) is ambiguous, which it is not, intuitively speaking.

   Roberts is aware that the power of her theory is a mixed blessing, and proposes a number of constraints on the notion of antecedent accommodation, which I want to discuss briefly. The constraints that Roberts suggests are the following: (i) modal subordination 'requires nonfactual mood' (1989: 701); (ii) 'it must be plausible that the modally subordinate

utterance has a hypothetical common ground suggested by the immediately preceding context' (ibid.); and finally (iii) modal subordination may not make available to anaphoric expressions antecedents that have no explicit representation in the given DRS (1989: 705).

Roberts's first constraint restricts modal subordination to utterances in non-factual mood. This constraint is needed to explain why the second *it* in (17) cannot link up to the indefinite *a bear*:

(17)     ?Fred may have seen a bear$_i$. It$_i$'s a big one.

This example shows quite clearly that, without any further constraints, the notion of antecedent accommodation nullifies the restrictions that DRT imposes on anaphoric reference. Pending rather severe restrictions, any theory which adopts antecedent accommodation is a null theory of anaphora, which predicts that anything goes, and the constraints Roberts proposes are intended to regain the effect of DRT's accessibility restriction.

The non-factuality constraint excludes the anaphoric link in (17) on the grounds that the second sentence is 'factual', which in this case correlates with the absence of an overt modal expression. However, Roberts's non-factuality constraint is weaker than it might appear to be at first sight, because she does not want to restrict modal subordination to sentences that contain overt modal operators. For example, she doesn't want this constraint to frustrate her analysis of (10), and postulates accordingly that on 'pragmatic grounds, we may assume that neither disjunct of a disjunction is asserted, and hence that both are nonfactual.' (1989: 702) It would seem, therefore, that this constraint is not nearly restrictive enough, because it allows modal subordination of any non-entailed clause, as in the following:

(18)     a.     ?Gurnemanz may have bought a new swordi. Kundry thinks
                 that iti is too heavy for him.

         b.     ?Fred may have seen a bear$_i$. It$_i$ can dance.

In these cases modal subordination does not seem to be possible. But as both examples obey the non-factuality constraint, Roberts's theory does not exclude this possibility. It would seem advisable, therefore, to strengthen this constraint by restricting modal subordination to sentences that contain overt modals, or that at least are interpreted as modals.

Roberts's second constraint requires that accommodated material be 'fresh' and hence must be copied from a recent portion of the discourse representation. It is this constraint, presumably, that applies to discourses like the following:

(19)     A thief might try to break in. So we should tell our neighbour that
         we're away for the weekend, and unchain the dog before we go out.
         If that doesn't scare *him* off, nothing will.

It is difficult to link up the italicised pronoun with the indefinite in the first sentence, and the difficulty is roughly proportional to the distance between the modally subordinated sentence and its 'antecedent'. The problem with this constraint is that it is too strict, as witness:

(20)     A thief might try to break in. So we should tell our neighbour that we're away for the weekend, and unchain the dog before we go out. If that doesn't scare *the intruder* off, nothing will.

Here I have merely replaced the pronoun with a full definite NP, and the resulting discourse is much better than the one in (19). The constraint Roberts proposes is plausible only so long as we confine our attention to anaphoric pronouns.

The same holds for Roberts's third constraint. It is well known that anaphoric expressions generally require explicit antecedent expressions: normally speaking it is not sufficient that a suitable antecedent can be inferred. The following minimal pair illustrates this contrast:

(21)     a.     Mary's head was chopped off but even so it kept smiling.

         b.     Mary was decapitated but even so it kept smiling.

In (21a) pronominal anaphora to Mary's head is possible, but in (21b) it is quite difficult, the intuitive reason being that in the second case the head in question has not been mentioned explicitly — it is merely inferred. However, this restriction only applies to descriptively attenuate presuppositions; the following is fully acceptable:

(22)     Mary was decapitated but even so her face kept smiling.

Roberts's constraint is designed to ensure that antecedent accommodation does not make accessible reference markers that don't correlate with a suitable explicit expression in the previous discourse. Consequently, this constraint not only rules out (23a), which is as it should be, but it blocks (23b), too, which is perfectly fine:

(23)     a.     We might decapitate Mary but it would keep smiling.

         b.     We might decapitate Mary but her face would keep smiling.

To sum up: Roberts's constraints on antecedent accommodation are deficient in a number of ways, and to the extent that they are adequate they require that we confine our attention to anaphora and ignore other presupposition inducers. But there is a more fundamental problem with these constraints: they are evidently ad hoc. All their flaws aside, these constraints seem plausible enough, but it shouldn't be necessary to stipulate them; they should be theorems. Something like Roberts's second constraint, for example, should follow from the descriptive paucity of anaphoric pronouns. But if we

adopt such a powerful antecedent generator as the one proposed by Roberts it is hard see how we can avoid having to stipulate these constraints.

## 6.2 Modals presuppose their domains

This slogan sums up the hypothesis that I want to consider in the remainder of this chapter. We have seen that the context dependence of modal domains, which is the key element in Roberts's theory, is widely recognized in the literature on modality. If my hypothesis is correct this context dependence can be captured as follows: modals are presuppositional expressions, and modal domains are presupposed. This hypothesis has repercussions for the construal of presuppositional expressions that occur in the scope of modals, in the following way. If a modal expression $\psi$ presupposes its domain, it makes sense to say that it can link up to another modal $\varphi$ via this presupposed domain, in the process of which entities that were introduced in $\varphi$ become accessible from within $\psi$. Thus the presupposition that a modal induces may make accessible discourse entities that would otherwise have remained inaccessible.

If this is correct, then there are no relevant differences between modal expressions and attitude verbs, and the theoretical machinery developed in the last chapter can be used to deal not only with attitude reports but with modal expressions as well. I will presently show that this is in fact in the case; but first I want to enlist independent evidence in favour for my hypothesis. If a modal presupposes its domain then modal domains should exhibit the projection behaviour that is characteristic of presuppositional inferences. And they do, as the following observations illustrate.

(24)    a.    It would have rained.

b.    It wouldn't have rained.

c.    It either would have rained or it would have snowed.

d.    If today had been Monday, it would have rained.

(24a) clearly requires a context to 'fill in' its domain. The intuitive meaning of this sentence is something like, 'Under such and such circumstances, it would have rained', and it depends on the context what the circumstances are. My hypothesis is that this context dependence is of a presuppositional nature, and this is corroborated by the remaining examples in (24); for (24b, c) exhibit exactly the same context dependence as (24a), while (24d) does not. This is precisely what we should expect if *would* presupposed its domain, so presumably that is what it does.

Assuming just that modals presuppose their domains, we can already account for two of the constraints on modal subordination that Roberts had to postulate (once the theory has been spelled out in more detail the third one will be derivable too). In the first place, it is no longer a mystery why modal subordination requires non-factual mood: this is simply because it requires a modal to trigger the presupposition that licenses modal subordination. Secondly, it also becomes clear why modal subordination prefers to pick up material from the immediately preceding context. In general, the restrictions that a modal expression imposes on its domain are not overly specific. The verb *would*, for example, merely presupposes that a non-factual context is given (i.e. under consideration), which is not very informative in comparison to the presuppositions induced by attitude verbs, for example. In this respect, modal expressions are not unlike anaphoric pronouns, and it is only to be expected that they will generally prefer to pick up 'antecedents' that are in focus.

Let us now consider in more detail how our hypothesis can be fleshed out. The main idea is that a modal expression $\alpha$ establishes a relation between a modal domain $\sigma$ and some $\sigma'$ which extends $\sigma$, and which I shall refer to as $\alpha$'s 'co-domain'; the latter corresponds with what is generally known as $\alpha$'s 'nuclear scope' or just its 'scope'. At DRS level domains and co-domains are represented by propositional reference markers, which are connected by relational constants. For concreteness, let us explicitly define two such constants, '$\square$' and '$\diamond$', as follows (as in the last chapter, $\sigma$ and $\sigma'$ are indexed propositions):

(25)    a.    $I_w(\square) = \{\langle \sigma, \sigma' \rangle \mid w(\sigma) = w(\sigma')\}$

        b.    $I_w(\diamond) = \{\langle \sigma, \sigma' \rangle \mid w(\sigma) \cap w(\sigma') \neq \varnothing\}$

This is in fact a straightforward transposition into our DRT framework of the classical possible worlds analysis of modality. It is possible to incorporate more sophisticated analyses, too, perhaps along the lines suggested by Frank (1997), but as I will argue in the following section, this would not affect our account of modal subordination in any fundamental way.

(26b) is the DRS that the grammar should associate with (26a):

(26)    a.    A thief might break into the house.

        b.    $[p, q: q = p+[x: \text{thief } x, \text{break-in } x], p \diamond q]$

The representation of the modal *might* involves two propositional reference markers, p and q, with p representing the modal domain, which is presupposed, and q the co-domain. Although the representation illustrated by (26b) will suffice to explain modal subordination, it is incomplete in some respects, of which I want to mention only one for the moment. According to (26b) there are no restrictions at all on the modal domain of *might*,

represented here by the reference marker p. It is quite obvious, however, that there are such restrictions (even if they are non-specific and vague, as I claimed two paragraphs ago), and that they play a role in the interpretation of modals. This point will be taken up in the following section, where it is shown that these restrictions are crucial if we want to get the details of modal subordination right. But I will ignore them for the time being.

The presupposition of the modal in (26a) could not normally be construed by means of accommodation, but let us suppose that it is accommodated nonetheless, so that we can go on with Roberts's example:

(27)    A thief might break into the house. He would take the silver.
        (= (6))

(28)    a.    $[p, q, p', q':$
              $q = p+[x: \text{thief } x, \text{break-in } x], p \diamondsuit q,$
              $q' = p'+ [\underline{z}: \text{take-silver } z], p' \square q']$

        b.    $[p, q, p', q': p' = q,$
              $q = p+[x: \text{thief } x, \text{break-in } x], p \diamondsuit q,$
              $q' = p'+[\underline{z}: \text{take-silver } z], p' \square q']$

        c.    $[p, q, p', q': p' = q,$
              $q = p+[x, z: z = x, \text{thief } x, \text{break-in } x], p \diamondsuit q,$
              $q' = p'+[: \text{take-silver } z], p' \square q']$

        d.    $[p, q, q':$
              $q = p+[x: \text{thief } x, \text{break-in } x], p \diamondsuit q,$
              $q' = q+[: \text{take-silver}(x)], q \square q']$

Once the first sentence has been dealt with, the DRS representation of (27) will be (28a). In (28a), there are two relevant presuppositions: the domain of *would*, represented by p', and the anaphoric presupposition triggered by the pronoun *he*. The presupposed domain can be bound to either p or q. The binding theory predicts that the latter option is preferred, because this allows the pronominal anaphor to be bound. For, once p' has been bound to q, as shown in (28b), x is accessible to z, and we can identify z with x so as to obtain (28c), which is equivalent with (28d). The meaning of this DRS is the following:

(29)    $[(28d)]_{\langle w, f \rangle} =$

        $\{\langle w, g \rangle: f \subseteq g \ \& \ \text{dom}(g) = \{p, q, q'\} \ \&$

            $g(q) = \{\langle v, j \rangle \mid \langle v, h \rangle \in g(p) \ \& \ h \subseteq j \ \& \ \text{dom}(j) = \{x\} \ \&$
            $j(x) \in I_v(\text{thief}) \cap I_v(\text{break-in})\} \ \&$

            $w(g(q)) \subseteq w(g(p)) \ \&$

$$g(q') = \{\langle v, j \rangle \in g(q) \mid j(x) \in I_v(\text{take-silver})\} \ \&$$

$$w(g(q)) = w(g(q'))\}$$

It should be noted that this is weaker than the intuitive interpretation of (27), because (28d) imposes no constraints on the propositional reference marker p, save for the fact that it is extended by q. Nonetheless, I maintain that (28d) is adequate as far as it goes; it is just that it is not a complete representation of the intended interpretation of (27). (28d) is incomplete for two reasons. To begin, I have thus far ignored the restrictions which the verb *might* imposes on its domain. If I hadn't done so, the semantic representation of (27) would have been more specific. It wouldn't have been much more specific, though, because the verb *might* does not impose very tight restrictions on its domain, as I pointed out earlier. The crucial constraint on the interpretation of (27) that I have left out of account is rather that this sentence would normally be uttered in a context in which the domain of *might* can be bound to some salient proposition. I have ignored this because my principal aim in this chapter is to account for modal subordination, and therefore I am mainly interested in the relation between the two sentences in (27). When I assumed that the domain of the first modal is accommodated, it was merely because I wanted to get it out of the way.

If we want to construct a more specific representation for (27), we need more information about the context in which this sentence is uttered. Imagine the following scenario. Fred and Wilma are on their way to the cinema, when the following conversation unfolds:

(30)    Wilma:  Fred?
   Fred:   Yes Dear?
   Wilma:  Did you check the back door?
   Fred:   No. Why?
   Wilma:  A thief might break into the house.

Suppose that the reason why Wilma asks Fred if he checked the back door is that she wants to know if the back door is locked. Fred's negative answer implies that it hasn't been established yet, at least as far as Wilma is concerned, that the back door is locked, so for all she knows the back door might not be locked, and after Fred's last turn in this miniature conversation, the proposition that the back door is unlocked is salient, and ready to be picked up to serve as the domain of the verb *might*.

I prefer not to view (30) as an instance of modal subordination because in this case linguistic and contextual factors interact in a way that is different from what one observes in the standard cases of modal subordination. According to the theory I have proposed, modal subordination occurs whenever a modal *expression* sets up an antecedent for another modal expression; this is what happens in (27), for example, where *might* sets up an

antecedent which is picked up by *would*. So we can account for the connection between the two sentences in (27) even if we abstract away from the context in which this discourse takes place. In (30) this is not possible, obviously. The difference between (27) and (30) may be likened to the difference between (occurrences of) anaphoric pronouns with and without linguistic antecedents. The verb *would* in (27) is similar to a pronoun which picks up a referent introduced previously by an indefinite NP, say; whereas the verb *might* in (30) is more like a pronoun that picks up a referent that has been made salient by other means.

In the case of pronouns it is usually obvious how their referents entered the discourse. In the case of modals it is not always so clear, and therefore the distinction between modals with and without linguistic antecedents may sometimes be hard to draw. But it is an important distinction nonetheless.

(31)     I don't have a microwave oven. I wouldn't know what to do with it.

Here it seems possible for the pronoun in the second sentence to link up to the indefinite in the first, and this case is similar to the prototypical instances of modal subordination in that the anaphoric link depends upon a modal expression: if we replace *wouldn't* by *don't* or suppress it altogether, anaphora is not possible anymore. It has been suggested by Roberts (1996) and Frank (1997), among others, that this is a case of modal subordination, too, and that, consequently, the same mechanism that is invoked to deal with (27) should deal with (31). I don't think that is correct. If (31) is just another instance of modal subordination, then, presumably, the negation in the first sentence sets up an antecedent for the modal in the second. So the negation sign must be given an intensional interpretation, along the following lines. Suppose that we extend our DRS language with conditions of the form '$p := \varphi$', whose interpretation specifies that the semantic value of $p$ is the proposition expressed by $\varphi$.[5] With the help of this construct, we could analyse the first sentence in (31) as follows:

(32)     [p: p := [x: microwave x, I own x], not p]

On this analysis, the first sentence in (31) might be paraphrased as follows: 'Apropos of the proposition that I own a microwave oven: it is not true.' Thus a negative sentence sets up a suitable discourse entity for a modal expression to pick up, and (31) can be treated as a case of modal subordination.

To see what is wrong with this proposal, consider the following examples:

(33)     a.     ?I haven't studied your proposal. I might support it.

---

[5]   It is not particularly difficult to develop this in more formal detail, but I will not do so here, because the main point I want to make does not hinge upon such details. See Frank (1997) for a more elaborate proposal along these lines.

b.  ?Nobody turned up at Betty's party. He would have seen her wearing a gorgeous evening dress.

c.  ?It's pitch dark in here, you can't see a thing. It might be a vampire bat.

If the proposed treatment of (31) were correct, these discourses should be fine (on their intended modal-subordination readings). For example, (33a) should mean something like, 'I haven't studied your proposal, but if I had, I might support it'. But as it is this discourse is simply infelicitous, and the same holds for (33b, c). This is an embarrassment for the intensional analysis of negation, which is aggravated by the following observations, which show that each of the examples in (33) can be matched by a genuine and felicitous instance of modal subordination:

(34)  a.  I would study your proposal, and I might support it.

b.  Someone might have turned up at Betty's party, and he would have seen her wearing a gorgeous evening dress.

c.  If it weren't so dark in here, you might see something, and it might be a vampire bat.

It has often been observed that negative sentences tend to evoke contexts in which the corresponding positive concepts are already salient (see Horn 1989 for an extensive overview). The first sentence of (31) would seem to be most likely to occur in a context in which 'having a microwave oven' is the topic of conversation. The intensional analysis of negation may be seen as an attempt to capture this observation in semantical terms, but the data I have mustered strongly suggest that this attempt is misguided. It appears that, when a negative sentence makes salient a positive concept, this is due to pragmatic processes, which have nothing to do with negation per se. (31) is not an instance of modal subordination.

I have shown how modal subordination can be accounted for in a way that is in line with our pre-theoretical intuitions about this phenomenon, but without the help of a copying mechanism. At the same time, the theory is considerably more restrictive than Roberts's. It entails, for example, that the usual accessibility restrictions apply to the selection of modal domains. This prediction is corroborated by observations like the following:

(35)  a.  Fred can read and he can write, too.

b.  ?Perhaps Fred can read. He can write, too.

In (35a), the presupposition triggered by the focus particle can be bound to an antecedent, but in (35b) a binding interpretation is not possible. Given that the presupposition triggered by *too* is difficult to accommodate

(cf. § 2.5), the binding theory predicts the contrast between these sentences. Roberts's theory, on the other hand, would allow the presupposition to be bound in both cases, and therefore it doesn't account for the unacceptability of (35b).

## 6.3  Constraints on modal subordination

The semantics of the modal predicates '◇' and '□' as defined in (25) is essentially the classical one, but our account of modal subordination doesn't stand or fall with this interpretation. We have seen that the domain of a modal expression is determined by contextual factors, and have argued that this context dependence is of a presuppositional nature: modals presuppose their domains. Kratzer (1981, 1991) observes, however, that modals are context sensitive in a further respect. Namely, the interpretation of a modal may be dependent upon the internal structure of its domain. Kratzer offers several arguments why this should be so, but the easiest and probably most convincing way of appreciating this point is by noting that modality is very often a matter of degree: a proposition can be more or less likely, a course of action more or less desirable, and so on. A sentence of the form 'It is likely that φ', for example, doesn't merely say that φ is the case in some world that belongs to a given domain of epistemic alternatives. What this sentence conveys is something stronger: φ must be the case in those worlds that represent comparatively likely alternatives. In general, therefore, modals are doubly relative: to a domain and to what Kratzer calls an 'ordering source', i.e. a partial ordering of the domain under consideration. A possible world may be closer to or farther away from what the Law provides, what the Lord ordains, what makes One feel good, and so on.

   If we decide to adopt Kratzer's amendment to the classical theory of modality, the interpretation of a modal expression comes to depend on three parameters: modal domain, quantificational force, and ordering source. But although it may be that a more adequate theory of modality requires each of these, this extension doesn't affect our account of modal subordination in any way, because this account leads us to expect that modal subordination is contingent upon the first parameter only: quantificational force and ordering source don't play a role at all. For example, in order for modal subordination to occur and enable anaphora from z to x in the following DRS,

(36)      $[p, q, p', q': q = p+[x: ... ], q' = p'+[z: ... ]]$

the propositional reference marker $p'$ must be bound to q, and the only prerequisite for this is that there be a match between the contents of $p'$ and

q. Thus our theory implies that there is no *a priori* reason to assume that modal subordination is contingent either upon the force of the modal expressions involved (for example, existential vs. universal) or upon their respective ordering sources. The only requirement is that their (co-)domains be compatible. Therefore, even modals of different types (for example, deontic and epistemic) should be able to engage in modal subordination, provided their domains match. The following examples corroborate these predictions:

(37)  a.  You must carry an umbrella but you may leave it at the reception.
      b.  Harry may carry an umbrella but he must leave it at the reception.

The modals in (37a) have universal and existential force, respectively, while in (37b) it is the other way round, and since modal subordination is possible in either case, these examples confirm that the force of the modals involved doesn't matter. Moreover, whereas most hearers would probably prefer to construe both modals in (37a) deontically, it is clearly possible to interpret the first modal in (37b) epistemically and the second deontically (although a deontic-deontic reading is possible too), without thereby precluding the possibility of modal subordination. Since it is clear that epistemic and deontic modals involve different kinds of ordering sources, this example corroborates the second prediction.

I predict that the possibility or impossibility of modal subordination turns upon the constraints that modals impose upon their domains, but I have not said anything yet about what kind of constraints might be involved here. So what I have provided so far, strictly speaking, is merely a framework for analysing the phenomenon of modal subordination. A comprehensive account of modal subordination would require an in-depth analysis of all the various types of modal expressions. This is a tall order, obviously, and it would take me far beyond the aims of this book. So let me just discuss a couple of examples.

As we proceed it will become clear that in analysing the restrictions on modal subordination, tense and time are key ingredients, and I shall therefore assume that our semantics can be extended so as to account for temporal reference and quantification. Since it is immaterial to our purposes what the details of such a treatment of time looks like, I shall simply assume that a state is now a triple consisting of a world, a time point, and an embedding function, and that the class of models is suitably restricted. As far as I can tell, however, nothing hinges upon our adopting any particular model of time.

Let us now take a closer look at the examples in (37). We said that in (37a) both *must* and *may* are most likely to be read deontically. Why is modal

subordination possible in this case? Presumably, the domain of a deontic modal is a subset of the set of possible futures, and therefore the domains of the two modals in (37a) must be contained in the set of possible future states. But as a matter of fact the domains of *must* and *may* are more narrowly circumscribed than that. (37a) obviously requires a context in which some hypothetical future course of action on the part of the addressee is already under consideration. For example, the interlocutors may be involved in some doubtful transaction and be making arrangements for a conspiratorial meeting. In such a context the set of relevant futures will be narrowed down considerably, and it is such a set that the first modal in (37a) will have as its domain; let σ be this set. Then the co-domain of *must* will be that extension σ′ of σ in which the addressee carries an umbrella, and the first conjunct may be paraphrased as follows: 'If $s \in \sigma$ and $s$ is a state which approximates as closely as possible the speaker's preferences as to how the said meeting should proceed, then $s \in \sigma'$, too.' Now the second modal in (37a) presupposes the same kind of domain as the first one does, and the co-domain introduced by the first modal, i.e. σ′, is of the appropriate type: it is a contextually given set of possible future states. Therefore, the modal *may* can select σ′ as its antecedent, as a result of which the indefinite becomes a possible antecedent for the pronoun *it*.

The proposed account of (37b) mirrors that of (37a), even if the first modal is read epistemically and the second deontically. The domain of the epistemic *may* picks up some contextually given subset of the possible futures, which its co-domain restricts to those states in which Harry carries an umbrella, and the deontic *must* can pick up this co-domain because it is precisely the kind of thing that it presupposes. Again, the explanation is couched entirely in terms of the modals' domains and co-domains; quantificational force and ordering sources don't enter the picture at all.

In these analyses I have emphasized the role of temporal constraints, because I believe that these are crucial to modal subordination. The following variations on (37a, b) illustrate this:

(38)    a.    ?You must carry an umbrella but you may have left it at the reception.

        b.    ?Harry may carry an umbrella but he must have left it at the reception.

These are both infelicitous (on their intended readings), and it is intuitively obvious why: in both cases the first conjunct refers to the future while the second one is about the past. The following example is slightly more demanding, but here temporal factors are essential, too:

(39)    ?Fred can balance a banana on the tip of his nose. He may eat it.

The discourse in (39) is awkward, and this has to do with the types of modality that can and may allow for, and in particular with their presuppositional requirements. As before, may in (39) can be read either epistemically or deontically, which doesn't however make a relevant difference to its presuppositional requirements. The modal can in (39), on the other hand, is most likely to get an 'ability' interpretation: the first sentence in (39) prefers a reading on which it states that Fred masters a certain skill. This ability interpretation can be characterized as follows. Let the domain of can be a set s of states that agree with the current state in that Fred has, say, the same acrobatic skills that he has here and now; its co-domain is that extension s¢ of s in which Fred balances a banana on the tip of his nose. What the first sentence in (39) says, then, is that s¢ is not empty.

Intuitively, there are two ways in which one might view the temporal connection between the domain of an ability modal and the utterance time. On the one hand, one might say that the temporal dimension is irrelevant to the interpretation of ability modals. On this type of account ability modals are basically timeless, and the domain of *can* in (39), that is σ, contains all states in which Fred has the same skills that he has here and now, regardless of whether these states lie in the past, present, or future. On the other hand, one might hold that the states in σ must be contemporaneous with the current state. Both alternatives seem plausible enough, but fortunately there is no need to choose between them here, for either choice allows us to explain the data. Suppose that we opt for the first alternative; σ will then be some set of states whose temporal index is arbitrary, and the same will hold for σ′, which is the co-domain of *can*. But since the second modal requires as its domain a given set of *future* states, the co-domain of *can* is not an appropriate antecedent, and modal subordination is blocked. The same conclusion follows if we take σ to be a set of contemporaneous states. So in either case we obtain an explanation for the awkwardness of (39).

To sum up, I have presented a number of cases in which modal subordination was possible and some in which it was not, and I have argued that these facts could be explained within the framework outlined in the previous section. In each case, an explanation could be given in terms of what modals presuppose (their domains) and the objects that they introduce (their co-domains). Every modal — more accurately: every *occurrence* of a modal — imposes certain restrictions on its domain and co-domain, and it is these restrictions that decide if modal subordination is possible or not. My claim is that what I have done here for a handful of examples can be done for sequences of modals in general, and if that is correct, we don't need a special theory of modal subordination. The facts about modal subordination are accounted for by the binding theory in conjunction with an adequate semantics of modal expressions.

## 6.4 Conditionals

In his paper on quantificational adverbs, Lewis (1975) argues that in examples like the following,

(40)    a.    If a farmer owns a donkey, he $\left\{ \begin{array}{c} \text{always} \\ \text{sometimes} \\ \text{usually} \end{array} \right\}$ beats it.

   b.    $\left\{ \begin{array}{c} \text{Always} \\ \text{Sometimes} \\ \text{Usually} \end{array} \right\}$ , if a farmer owns a donkey, he beats it.

the *if*-clause serves to restrict the domain of the adverb, which he analyses as an unselective quantifier. Lewis observes that there may be any number of *if*-clauses in constructions like these, and that the number may be zero, in which case (40a) and (40b) collapse into:

(41)    A farmer $\left\{ \begin{array}{c} \text{always} \\ \text{sometimes} \\ \text{usually} \end{array} \right\}$ beats his donkey.

Here the domain which the adverb quantifies over remains implicit. On the other hand, Lewis notes, ordinary conditionals may be construed as containing a covert adverb of quantification. For example,

(42)    If a farmer owns a donkey, he beats it.

is to be treated just like the examples in (40), but with an implicit *always*. It is this view on conditionals which lies at the heart of DRT.

   Building on Lewis's work, Kratzer (1979, 1981, 1991) proposes a related analysis of conditionals.

(43)    If a murder occurs, the jurors must convene. (Kratzer 1991: 648)

Instead of analysing (43) in terms of a binary connective, Kratzer proposes that the *if*-clause serves to restrict the domain of *must*. Thus she construes (43), in effect, as 'The jurors must convene', on the understanding that the modal domain is restricted to worlds in which the antecedent of (43) is true. Kratzer extends this analysis to ordinary conditionals without overt modals, like:

(44)    If a murder occurs, the jurors convene. (Kratzer 1991: 649)

which she claims contains a covert necessity operator. The resulting theory is a generalization of the theory of conditionals pioneered by Stalnaker (1968, 1975) and Lewis (1973).

On the Kratzer/Lewis analysis, the function of an *if*-clause is to restrict to the domain of a quantifier over cases, which may or may not be overtly realized.[6] If the quantifier is left implicit, it will be supposed by default that it has universal force, and I will follow Kratzer in assuming that such implicit quantifiers are generally modal ones.

(45)    a.    If Mary had a car, she would take me to work in it. I could drive the car too. (Karttunen 1976: 375)

        b.    If Jack has sisters, Jack's sisters have insisted his wife have children. If Jack's sisters have insisted his wife have children, Jack's wife has let herself have children. If Jack's wife has let herself have children, Jack's wife has become neurotic. If Jack's wife has become neurotic, all of Jack's children are bald. (Kuroda 1979: 185)

In (45a) the antecedent of the conditional introduces a car which is taken up within the scope of the modal *could*. In (45b) a peculiar chaining effect occurs because each *if*-clause picks up where the preceding conditional left off. Thus, in a sense, the whole of the preceding discourse is taken up by the *if*-clause of the last conditional, as a result of which the noun phrase *Jack's children* in the last sentence gets access to its intended antecedent.

If we adopt the Kratzer/Lewis analysis of *if*-clauses, it is obvious, at least in outline, that these observations mesh rather nicely with our account of modal subordination. For if we incorporate the Kratzer/Lewis analysis into the theory of modals presented in the foregoing, it follows immediately that *if*-clauses are presuppositional expressions. This is not to say, of course, that a conditional presupposes the truth of its antecedent. It is to say, rather, that a conditional triggers the same kind of presupposition as a modal expression, the main difference being that the *if*-clause imposes overt restrictions on the content of this presupposition. (It follows from this that the presuppositions triggered by conditionals are richer than the presuppositions triggered by modals, and will therefore be easier to accommodate; and so it appears to be.)

In the preceding section, we already talked about restrictions on the presuppositions triggered by modal expressions, but we didn't show how such restrictions can be represented. This is not such a difficult task, however. In general, the problem is to impose upon a given reference marker $p$ the restriction that $p$ satisfies a set of restrictions represented by a DRS $\varphi$. In other words, $\varphi$ may not contain information that is not already contained in $p$, which is to say that $p = p+\varphi$. If this is the case, we say that $p$'supports' $\varphi$, and write '$p \vdash \varphi$':

---

[6]  The first one to see this connection between *if*-clauses and domain restriction may have been Belnap (1970), who is mentioned by Lewis in this connection.

(46)      $p \vdash \varphi =_{def} p = p + \varphi$

As a general format for representing conditionals of the form *if* S *then* (A) S′, where A is some modal expression, I now propose the following:

(47)      $[p, q: p \vdash \varphi, q = p + \varphi', p \ R_A \ q]$

Where $\varphi$ and $\varphi'$ are the DRSs corresponding to S and S′, respectively, and $R_A$ is a predicate correlated with the modal expression A. This schema retains the dynamic character of the standard DRT conditional, albeit in a somewhat roundabout way: since $p$ supports $\varphi$, all reference markers in $\varphi$ must be defined in the semantic value of $p$, and are therefore accessible from $\varphi'$. In accordance with Kratzer's view, I assume that, if A is empty, $R_A$ defaults to some form of necessity.

To illustrate this proposal, the last two sentences of Kuroda's sample discourse in (45b), would be represented as in (48a):

(48)      a.     $[\underline{p}, q, \underline{p}', q':$
                       $p \vdash$ [x: x are children of J's wife],
                       $p \ \square \ q,$
                       $q = p + $[J's wife gets neurotic],
                       $p' \vdash$ [J's wife gets neurotic],
                       $p' \ \square \ q',$
                       $q' = p' + [\underline{z}: \underline{\text{z are J's children}}, \text{bald z}]]$

         b.     $[p, q, p', q': p' = q,$
                       $p \vdash$ [x, z: z = x, z are J's children, x are children of J's wife],
                       $p \ \square \ q,$
                       $q = p + $[J's wife gets neurotic],
                       $p' \vdash$ [J's wife gets neurotic],
                       $p' \ \square \ q',$
                       $q' = p' + [: \text{bald z}]]$

         c.     $[p, q, q':$
                       $p \vdash$ [x: x are J's children, x are children of J's wife],
                       $p \ \square \ q,$
                       $q = p + $[J's wife gets neurotic],
                       $q \ \square \ q',$
                       $q' = q + [: \text{bald x}]]$

The somewhat schematic DRS in (48a) contains three presuppositions, two corresponding to the *if*-clauses of the second pair of conditionals in (45b), and one triggered by the definite noun phrase *Jack's children*. Once this representation is in place, the projection mechanism of the binding theory can take its by now familiar course. The first presupposition, p, cannot be

bound and is therefore accommodated.[7] Then p′, which represents the *if*-clause of the second conditional, is bound to q, as a result of which x becomes accessible to z, and the presupposition triggered by *Jack's children* can be resolved. The outcome is (48b), which is equivalent to (48c).

If *if*-clauses are presupposed, as I suggest, we also obtain an explanation for examples like (49a):

(49)    If we get a taxi, and if we aren't held up in a traffic jam, and if there isn't an air strike, then we'll be in Timbuktu before teatime.

Observe that the second and the third *if* in (49) are redundant; they can be eliminated *salva veritate*. The presuppositional analysis of *if*-clauses explains why this should be so: the first *if*-clause sets up a propositional reference marker which is passed on to the consequent by the subsequent *if*-clauses, and whose range of possible values is narrowed down in the process.

This presuppositional analysis of *if*-clauses accords with Haiman's (1978, 1986) proposal that *if*-clauses are *topics*. Haiman notes that there are several parallels between topic phrases and *if*-clauses. For example, the antecedent of a conditional normally precedes the consequent,[8] and in many languages *if*-clauses carry topic markers. Another parallel, which Haiman does not mention, is that *if*-clauses can be topicalized:[9]

(50)    a.    If the accused is guilty then I suppose everybody hopes he will be hanged.

        b.    I suppose everybody hopes the accused will be hanged if he is guilty.

(50a) can be construed in such a way that it is synonymous with (50b), and on this reading the *if*-clause has presumably been fronted from a position within the scope of the quantifier. These observations lend support to Haiman's claim that *if*-clauses are topics, and given the family ties between topics and presuppositions they indirectly support the presuppositional analysis, too.

A further piece of evidence in favour of this analysis is the well-known fact that, if people are arguing about whether 'if φ then ψ' is true or not, they will assume φ as given, and focus their argument on ψ. As Ramsey puts it in a famous footnote:

> If two people are arguing about 'if p, will q?', and are both in doubt about p, they are adding p hypothetically to their stock of knowledge, and arguing on that basis about q. (Ramsey 1931: 247)

---

[7]  Of course, this presupposition would actually be restricted by the preceding discourse; so strictly speaking the interpretation we are deriving will turn out to be too weak, or, rather, incomplete (cf. § 6.3).

[8]  This is Greenberg's (1963: 84–85) Universal of Word Order 14. Cf. also Comrie (1986: 83–86).

[9]  I owe this observation to Paul Kay.

There has been some discussion in the philosophical literature whether the principle of conditional excluded middle is valid or not:

(51)    (if $p$, then $q$) or (if $p$, then not-$q$)

Stalnaker (1981) argues that this principle is a valid one; Lewis's (1973) analysis of conditionals implies that it isn't. Consider the following pair of examples, devised by Quine (1950: 14):

(52)    a.    If Bizet and Verdi had been compatriots, Bizet would have been Italian.

        b.    If Bizet and Verdi had been compatriots, Verdi would have been French.

While on Stalnaker's analysis both conditionals would lack a truth value, on Lewis's account they would both be false. Although I have a slight preference for Stalnaker's position, I have no particularly strong opinion one way or the other, and mainly want to note here that the analysis of conditionals I propose throws a new light on this issue. If *if*-clauses are presupposed, then the question whether (51) is valid is tied up with the problem of presupposition failure. It may be argued that the presupposition triggered by an *if*-clause may fail, for instance because the proposition it presents as given is insufficiently specific. If this is the case, we are in the empirical quicksand briefly surveyed in § 1.6. This is doesn't improve the situation, to be sure, but at least it gives us a better idea where we are.

It is an almost stale observation that there is a close connection between *if*-clauses and sentences beginning with 'Suppose that . . .', 'Assume that . . .', 'Given that . . .', and so on. This connection is noted by Kamp (1988), too, but he concedes that his version of DRT does not explain it (cf. also Kamp and Reyle 1993: 141–146). On the presuppositional analysis of conditionals, by contrast, an explanation readily suggests itself.

(53)    a.    Suppose Fred is a Freemason. Then Barney is a Freemason, too.

        b.    If Fred is a Freemason, (then) Barney is a Freemason, too.

Stylistic nuances aside, (43a) and (43b) are equivalent. In the first half of (43a), the speaker invites his audience to suppose the world to be such that Fred is a Freemason, and in the second half he says that in such a world Barney is a Freemason, too. Interestingly, (43b) can be paraphrased in exactly the same way.[10] To explain this convergence, we only need to assume

---

[10] Note, incidentally, that (49) is matched by:

> Suppose we get a taxi, and suppose we aren't held up in a traffic jam, and suppose there isn't an air strike; then we'll be in Timbuktu before teatime.

See Kamp (1988) for further discussion of this sort of example.

that the first sentence of (43a) instructs the hearer to set up a reference marker for the proposition that Fred is a Freemason, as shown in (54a).

(54)   a.   [p: p ⊢ [F is a Freemason]]

       b.   [p, q, q′: p ⊢ [F is a Freemason], q □ q′,
            q′ = q+[B is a Freemason]]

       c.   [p, q′: p ⊢ [F is a Freemason], p □ q′,
            q′ = p+[B is a Freemason]]

Once the semantic representation of the second sentence has been added, we get (54b). This DRS contains only one relevant presupposition, which is triggered by the adverbial *then*. The presupposition is bound to the antecedent set up by the preceding sentence, and the interpretation we finally obtain is (equivalent to) (54c). This is the same as the interpretation we would get for (43b).

If this explanation is correct, then presumably natural-language conditionals are grammaticalized mini-discourses, which first present a hypothetical state of affairs, and then make a statement about it. On this account, conditionals derive from a source that is *doubly* presuppositional: the antecedent is presupposed, as I have argued already, and the consequent is anchored to the antecedent by means of a presuppositional bridge, which in English (like in many other languages) may be signalled by the particle *then*, whose pronominal origin can hardly be overlooked. In *if . . . then* conditionals this pragmatic bridge between antecedent and consequent has set into a grammatical connection.

### 6.5 Summing up

What Roberts has dubbed 'modal subordination' is actually a sub-pattern of presupposition projection that may be schematized as follows:

$$(\alpha \underline{\quad\quad}) \quad (\beta \underline{\quad\quad})$$

Here α and β stand for modal expressions, and modal subordination occurs whenever β's domain is bound to α's co-domain; in such cases presuppositions triggered within the scope of β can be bound within the scope of α. We have already seen that this pattern is not confined to modal expressions, as the same may happen with conditionals and attitude verbs. And the schema above is even more general than that, because it also holds for ordinary quantifying expressions. The binding theory gives a uniform account of all phenomena that instantiate this schema. For modals, conditionals, and attitude verbs, this was shown in the present chapter and the preceding one; the theory has been applied to extensional quantifiers by Geurts (1996c) and Geurts and van der Sandt (1999).

By way of summing-up I conclude this chapter with a list of examples which instantiate the schema above; they are my exhibit A. Unless indicated otherwise, the following examples are from Karttunen's landmark paper, 'Discourse referents', which had been haunting the linguistic community for seven years already when it appeared as Karttunen (1976):

*Attitude verbs:*

(55)   a.   Bill says he saw a lion on the street. He claims the lion had escaped from the zoo.

        b.   Tom expects to catch a fish and intends to fry it for dinner. (McCawley 1981)

*Modals:*

(56)   a.   You must write a letter to your parents. It has to be sent by airmail. The letter must get there by tomorrow.

        b.   It's certain that Sam will find a girl and possible that he will kiss her. (Lakoff 1972)

*Conditionals:*

(57)   a.   If Jack has sisters, Jack's sisters have insisted his wife have children. If Jack's sisters have insisted his wife have children, Jack's wife has let herself have children. If Jack's wife has let herself have children, Jack's wife has become neurotic. If Jack's wife has become neurotic, all of Jack's children are bald. (Kuroda 1979)

        b.   Ordinarily if it is raining, if $x$ is driving and sees $y$ walking, and if $y$ is $x$'s friend, $x$ offers $y$ a ride. (Lewis 1975)

*Quantifiers:*

(58)   a.   Harvey courts a girl at every convention. She always comes to the banquet with him. The girl is usually very pretty.

        b.   Every time Bill comes here, he picks up a book and wants to borrow it. I never let him take the book.

*Mixed cases:*

(59)   a.   Mary wants to marry a rich man. He must be a banker.

        b.   I wish Mary had a car. She would take me to work in it. I could drive the car too.

        c.   If Mary had a car, she would take me to work in it. I could drive the car too.

CHAPTER 7

# Names

What could be simpler than a name? An individual is given a name, and henceforth the name refers to the individual, and can be used to talk about it. It would seem that that is about it. Probably the main reason why Kripke's (1980) theory of names has been so influential is that, according to Kripke, this folk theory of names is essentially correct. On Kripke's account names are rigid designators; a name refers to the same individual in all possible worlds. A person is baptized *Aristotle*, say, and henceforth the name *Aristotle* refers to that person, irrespective of the context in which it occurs. Whether I say that Aristotle was wise, or that Aristotle might have become a carpenter, or that John believes that Aristotle *was* a carpenter — the name refers to Aristotle every time. If this is correct, then names are quite unlike definite NPs, and it may be argued that in this respect, too, Kripke's theory is in accordance with speakers' intuitions. Apparently, a definite NP such as *the tallest woman in the world,* for example, is not used like a name at all, and it certainly doesn't refer to the same person in all possible worlds.

Nevertheless, the notion that names are definite NPs has always had its advocates, although their number may have plummeted in the wake of Kripke's campaign against the description theory of names, as he dubbed it. There is one particular version of the description theory that I find appealing; it is the theory first proposed by Kneale (1962). According to this quotation theory, as I will call it, names are definite NPs, and the meaning of a name N is 'the individual named N'. At the end of the first lecture of *Naming and Necessity*, Kripke gives short shrift to this version of the description theory of names. I have always felt that Kripke's criticism of this view falls wide of the mark, and that the quotation theory is essentially correct. In this chapter I try to justify this assessment.

The purpose of this chapter is twofold. First, I want to present and defend a presuppositional version of the quotation theory, and show that the binding theory yields an account of names that is superior to Kripke's, not only because its predictions are much better but also because Kripke's theory is decidedly *ad hoc* in comparison. While I set out my argument, I will have the opportunity to further discuss certain aspects of the binding theory and its DRT framework; this is the second purpose of this chapter.

In the first half of this chapter I will argue for the quotation theory and against Kripke's proposal, without making very specific assumptions about the former. So the conclusions that I draw on the basis of this discussion do not stand or fall with the particular version of the quotation theory that I advocate, which is presented in the second half of the chapter. To begin with, I draw up a rather long inventory of observations whose purpose is to show that there are no empirical differences between names and overt definite NPs (§ 7.1). This inventory not only lends support to the quotation theory, it also comprises various sorts of counterexamples against the rigid-designator analysis. In the second section I discuss Kripke's chief argument against the quotation theory, and show how it can be dismantled (§ 7.2). In § 7.3, I merge the quotation theory with the presuppositional treatment of definites that I have been using in foregoing chapters, and explain how the resulting theory accounts for the data listed in § 7.1. In § 7.4, I try to account for the intuition of rigidity that informs Kripke's theory, and makes it seem so plausible. I close with some remarks for warding off misunderstandings, and a brief summary of why I believe Kneale was right, after all.

## 7.1 Names and other definite descriptions

The quotation theory claims that a name is synonymous with a definite NP of the form 'the individual named so-and-so'. For the time being we can leave the proposed analysis in this somewhat unspecific state, because I first want to raise a number of points that don't require a more precise formulation.[1] If the quotation theory is correct, then names must be expected to be used and interpreted like other definite NPs. If, on the other hand, names are rigid designators, then we should expect significant empirical differences between names and definite NPs. I will now show that it is the former prediction rather than the latter that is borne out by the facts.

The main objectives of this section are to take stock of the semantic properties of names and definite NPs, and to show that there aren't any fundamental differences between these two types of expressions. I will freely help myself to whatever technical jargon I find convenient for these purposes. For example, I will sometimes pretend as if names and definites

---

[1]  Although occasionally Russell came quite close to defending the quotation theory, to the best of my knowledge its first proponent was Kneale (1962), who was taken to task for this by Kripke (1980). Subsequently the theory received support from Loar (1976), Bach (1981, 1987), Cresswell (1985), and Fodor (1987). More recently, a number of presupposition theorists have swelled the ranks of the quotation theory: it is more or less taken for granted by van der Sandt and Geurts (1991), van der Sandt (1992), and Beaver (1993), among others. Zeevat (1996) argues at some greater length for a presuppositional version of the quotation theory that is more or less the same as mine.

had scope but will also talk of their 'referents'. This terminological *bric à brac* should not be taken too seriously.

Taken together, the following observations strongly suggest that names pattern with (other) definite NPs in all relevant respects:

• Names often take the form of definite NPs: *the United Nations, the Goldberg Variations, the Netherlands, the Annunciation, the Holy Spirit, the Bank of England*. These are names, no doubt, but they certainly look like definite NPs. In English, river names always carry a definite article (*the Danube*). In Italian, names of women often have, and sometimes must have, a definite article: *la Loren, la Carolina*. I suspect that all languages which have definiteness markers allow them to occur on proper names, and I wouldn't be surprised if there were languages in which names were always marked for definiteness.[2]

• So-called definiteness effects apply to names and definite NPs alike. On the one hand, if a construction selects for indefinite (or weak) NPs, names as well as definite NPs are excluded. This holds for English *there*-sentences, for example:

(1)     There is $\left\{ \begin{array}{c} \text{*John} \\ \text{*the philosopher} \\ \text{a philosopher} \end{array} \right\}$ available.

On the other hand, if a construction selects for definite NPs, it will accept names, too, as the case of the partitive construction illustrates:

(2)     half of $\left\{ \begin{array}{c} \text{Belgium} \\ \text{the country} \\ \text{*some countries} \end{array} \right\}$

• As Kripke (1980) points out, Donnellan's referential/attributive distinction applies to names just as it applies to definite NPs. In Kripke's example, two interlocutors observe Smith at a distance and take him to be Jones. Accordingly, they use the name *Jones* to refer to Smith (although Kripke insists that semantically speaking they are referring to Jones).

• Names can be used literally as well as non-literally (this point is clearly related to the previous one, but it is not exactly the same). For example, a man who jocosely refers to his wife as *the Queen* may use the name *Cleopatra* to much the same effect (cf. Bach 1987).

---

[2]  Strawson (1950) refers to definites like *the Glorious Revolution* and *the Great War* as 'quasi-names' or 'embryonic names', and remarks that an expression may 'for obvious reasons, pass into, and out of, this class (e.g. 'the Great War').' Although this observation is undoubtedly correct, I prefer to express it differently. Instead of postulating a separate class of 'quasi-names', I would say that the distinction between names and other definites is vague, because the relation 'being named N' does not have a precise meaning (cf. § 7.2).

• Names are like overt definites in that they can be used generically:

(3)          $\left\{\begin{array}{c}\text{The light bulb}\\\text{Coca Cola}\end{array}\right\}$ was invented by an American.

(4)          $\left\{\begin{array}{c}\text{The panzercroc}\\\text{Pristichampsus}\end{array}\right\}$ hunted during the Eocene Epoch, about 49
million years ago, but it was very rare. (Bakker 1988: 74)

Obviously, in these examples the names *Coca Cola* and *Pristichampsus* are used generically. It is true that it is difficult to think of a plausible scenario in which a proper name like *John* will have a generic interpretation, but this has nothing to do with the fact that *John* is a name. It is just that the kind that a generic occurrence of *John* would denote is of such limited use. It is precisely in this respect that *Coca Cola* and *Pristichampsus* are different from *John*, and this is why the examples above are felicitous.

• Definite NPs and names are typically, though not always, used to refer to objects that are part of the common ground between speaker and hearer. We may distinguish two cases here. On the one hand, definites and names may both be used to refer to objects that have been explicitly introduced into the discourse:

(5)          John has a poodle named 'Horace'. $\left\{\begin{array}{c}\text{Horace}\\\text{My poodle}\end{array}\right\}$ is three years
old.

On the other hand, names and definites may also be used to refer to individuals that are given in the larger situation in which the discourse takes place. For example, in some contexts either *the girl from no. 21* or *Julia* can be used to refer to my daughter, and in England *the Queen* and *Elizabeth II* will typically refer to one and the same person.

• Definites and names both have bound-variable uses:

(6)          a.    If I can choose between a red one and a green one, I'll have the red one.

             b.    If a child is christened 'Bambi', then Disney will sue Bambi's parents.

The definite NP in the consequent of (6a) is bound to the indefinite in the antecedent, and the same applies for the second occurrence of *Bambi* in (6b). People have objected to this observation on the grounds that (6b) is not quite felicitous, and although I do believe that this sentence is fully acceptable as it stands, I have some sympathy with this judgment. But if (6b) is vaguely odd, it is because the word *Bambi* is repeated for no good reason:

why use this word when a possessive pronoun would have done just as well?[3] If this is correct, then we should be able to come up with better examples if we can somehow motivate the repetition of the name. And we can, for instance by increasing the distance between the name and its antecedent and/or by introducing competing antecedents:

(7)    a.    If a child is christened 'Bambi', and Disney Inc. hear about it, then they will sue Bambi's parents.

        b.    The name of a product is essential to its commercial success. For example, if you want to buy washing powder and are given the choice between 'Black', 'White', and 'Drab', you will choose White, won't you?

It seems to me that (7a) is better than (6b), and as far as I can tell there is nothing wrong with (7b).

The following examples are cases of binding, too, intuitively speaking:

(8)    a.    Every time we do our Beatles act, [the guy who impersonates] Ringo gets drunk afterwards.

        b.    Every time John goes to see a performance of Hamlet, he falls in love with [the actress who plays the part of] Ophelia.

(9)    a.    Perhaps Mary has a son named 'John' and perhaps $\left\{ \begin{array}{l} \text{her son} \\ \text{John} \end{array} \right\}$ is the thief.

        b.    Mary is under the illusion that she has a son named 'John' and she believes that $\left\{ \begin{array}{l} \text{her son} \\ \text{John} \end{array} \right\}$ is the thief.

In (8a, b) the definite NPs as well as the names can be interpreted non-referentially. On this interpretation, *Ringo* is more or less equivalent with 'whoever impersonates Ringo', and the same holds, mutatis mutandis, for *Ophelia*. In (9a, b) there is a preference for a non-referential interpretation, and intuitively *her son* or *John* in the second conjunct is 'bound' in the first.

---

[3]   This is a familiar situation in the presupposition literature. It is widely held, for example, that in a sentence like the following the presupposition that there is a king of France, which is triggered by the definite NP in the consequent of the conditional, is somehow intercepted in the antecedent:

    If France has a king, then the king of France is bald.

Of course, this sentence is slightly awkward, too, and for the same reason: there is no apparent reason why a full definite should be used instead of a pronoun.

• Although definite NPs and names normally refer to objects in the common ground, they can be used to introduce new objects:

(10)      My best friend is $\left\{ \begin{array}{c} \text{my sister} \\ \text{John} \end{array} \right\}$.

Both variants of (10) can be used in a situation in which the intended referent is new to the hearer.

• Names generally take wide scope. In this respect, too, they are like many other definite NPs. Compare:

(11)      a.    The Prime Minister could have been rich.

          b.    The man could have been rich.

          c.    John could have been rich.

(11a) is ambiguous: it can either mean that the person who happens to be Prime Minister could have been rich, or just that we might have had a rich Prime Minister. (11b) and (11c), by contrast, only allow for readings of the first type.

     The suggestion that, with respect to scope, names are unlike definites in that they always take wide scope is incorrect for two reasons. First, as we have just seen, only some definite NPs alternate relatively freely between wide-scope and narrow-scope interpretations. Secondly, and this is the last point on my list,

• Names *can* take narrow scope, too. For reasons to be discussed below, this may require a somewhat outlandish type of context, but it does happen. For one thing, there are the bound-variable uses of names mentioned above. For another, there is the case of Aaron Aardvark, as reported by Bach (1987: 146–147):

> The electoral process is under attack, and it is proposed, in light of recent results, that alphabetical order would be a better method of selection than the present one. Someone supposes that 'Aaron Aardvark' might be the winning name and says, 'If that procedure had been instituted, Ronald Reagan would still be doing TV commercials, and
>
> [(12)] Aaron Aardvark might have been president.'

Clearly, in this scenario the speaker need not believe that there is anybody for the name *Aaron Aardvark* to refer to, yet (12) isn't infelicitous in any way. Further examples of the same type are:

(13)      a.    In English, Leslie may be a man or a woman.

          b.    But John is always male.

Although it is always hard to prove that something doesn't exist, the foregoing observations do suggest quite strongly that there are no deep semantic or pragmatic differences between names and overt definites. Such distinctions as must be made set some definite NPs apart from others but they don't draw the line between names and overt definites. Furthermore, most of the observations I have mustered are difficult to reconcile with the hypothesis that names are rigid designators. Not to mince words: they falsify it. If names were rigid designators it would be worrying, at the least, that names can have definite articles, that they can be used attributively and non-literally, and can be used to introduce individuals that are new to the hearer, and it would be a mystery that names have bound-variable uses and may take narrow scope.

It may be objected that at least some of these observations are irrelevant because they involve 'metalinguistic' uses of names, the implication being that names are generally not used this way.[4] I think this is partly right and partly wrong. What is right about it is that some of the data I have listed are somewhat out of the ordinary. But that is precisely what one should expect if the quotation theory is correct. If it is true that the meaning of a name is 'the individual named N', then the content of a name is special in a way that will make it difficult to construct examples in which names act as bound variables, for example. But the same holds for some overt definites, too. Consider *the cosmos*. There can be no doubt that this is a definite NP, but due to its truly comprehensive meaning it will be difficult to construct sentences analogous to the ones above in which *the cosmos* acts as a bound variable. Difficult, but not impossible. And although such cases will be very 'special', they will surely count as evidence that *the cosmos* is an ordinary definite NP. Therefore, I see no reason to dismiss the data presented in (6)-(9) and (12)-(13) merely because these are special cases in some sense.

It is understandable that someone who takes the position that names are rigid designators will want to claim that my counterexamples are irrelevant because they involve deviant uses of proper names; this would be Kripke's position, presumably (cf. Kripke 1980: 62, n. 25). However, this is a lame defence unless it can be shown on independent grounds that the *semantic values* of, say, *Bambi* in (6b) or *Aaron Aardvark* in (12) are non-standard. It will not do to note that these sentences are special or require special contexts, and leave the matter at that. In the absence of such independent evidence, these data stand as clear counterexamples against Kripke's theory of names.

---

[4] It may seem easy to distinguish 'metalinguistic' from 'ordinary' uses, but in fact it isn't easy at all. See Geurts (1998a) for discussion.

## 7.2 Names and reference

Kripke charges that the quotation theory is circular because it actually presupposes what it must explain, viz. how names come to refer:

> Someone uses the name 'Socrates'. How are we supposed to know to whom he refers? By using the description which gives the sense of it. According to Kneale, the description is 'the man called "Socrates" '. [...] We ask, 'To whom does he refer by "Socrates"?' And then the answer is given, 'Well, he refers to the man to whom he refers.' If this were all there was to the meaning of a proper name, then no reference would get off the ground at all. (Kripke 1980: 70)

But what if the quotation theory didn't explain how names manage to refer? Suppose we analyse definite NPs along Russellian lines. (Just suppose. I don't endorse Russell's analysis but I want to show that Kripke's criticism can be refuted without presupposing the analysis of definite NPs that I favour.) Then the quotation theory entails that names are non-referential expressions. This may strike some as an intuitively repugnant conclusion, but Kripke himself has argued persuasively that such intuitions are not necessarily detrimental to a Russellian theory. Critics of Russell's theory of descriptions have often claimed that the theory is falsified by referential uses of definite NPs. Kripke (1977) argues, however, that this is not so, because a Russellian may hold that the theory of descriptions is a *semantic* theory and need therefore only be concerned with what is said (in Grice's sense); intuitions about referential uses of definite NPs pertain to what is meant by a speaker on a given occasion. Put otherwise, the idea is that although semantically speaking definite NPs are non-referential expressions, speakers may use them for conveying information about specific individuals. But if names are definite NPs, then the same observations should apply to them, too. Hence, Kripke's complaint that the quotation theory is not a theory of reference not only presupposes that any semantical theory of names should be a theory of reference, like his own; it also denies the quotation theorist an account of reference that Kripke recommends elsewhere.

If NPs of the form 'the individual named N' are to be treated in a Russellian framework, they must probably be viewed as 'incomplete' definites, i.e. as on a par with *the table, the child*, and so on.[5] As Strawson was perhaps the first to point out, it would seem that such expressions, which rarely if ever describe unique objects in the world yet can be used without

---

[5]  This terminology is tendentious because only in the context of Russell's theory of descriptions is it necessary to consider such NPs to be incomplete. I use the term here merely because it has gained some currency in the Russellian tradition.

apparent difficulties, cause an especially severe problem for Russell's theory of descriptions. Russell's followers have countered Strawson's objection in various ways. According to Bach (1987), for example, a speaker who produces a sentence containing an incomplete definite NP virtually never says what he means (I am again speaking in Gricean terms): strictly speaking such sentences are false, but if all goes well, as it usually does, they will nonetheless manage to convey information about specific individuals.

This may or may not strike one as a plausible strategy for dealing with incomplete NPs, but this is as it may be, since for the moment I merely want to note that the quotation theory is immune to Kripke's objection that it doesn't explain how names come to refer. This objection clearly presupposes that names are referential expressions, where reference is to be understood as a semantic relation between a term and a real-world individual, and this presupposition cannot be taken for granted. Of course, it can hardly be denied that names may be used to convey information about real-world individuals, but the claim that names are referential expressions does much more than restate this banal truth in different terms, and should not be accepted uncritically.

Kripke's criticism is misguided for a further reason, as well, because it doesn't distinguish between two properties that should be strictly kept apart, viz. being the referent of N, and being named N.[6] That N is the name, or one of the names, of an individual *a* does not entail that N refers to *a* or that *a* is among N's referents. Kripke assumes as a matter of course that when we baptize an individual N we *eo ipso* determine the reference of the name N. This is not so. Beethoven's Missa Solemnis is commonly referred to as THE *Missa Solemnis*, but its name is *Missa Solemnis*. Similarly, an Italian may refer to Maria Callas with LA *Callas*, but her name was *Callas*.[7] The difference is slight (a mere two or three letters, after all) but telling: it highlights the fact that although it so happens that in English names can generally be used to refer, it might have been otherwise (and it sometimes is, even in English). The grammar of English might have dictated that in order to refer with the help of a name, it must always be preceded by the definite article, or, for that matter, by *the individual named*.

Bearing a name is like wearing a tie. Like ties, names are seldom unique, but circumstances permitting they may be used for referential purposes. More accurately, just as you can employ the attribute of wearing a tie to identify to your audience the person you have in mind (John, as the case may be), you can use the attribute of being named *John* for the same

---

[6]  This distinction is emphasized by Bach (1987), who proposes to speak of 'the bearer of N' instead of 'the individual named N', in order to forestall any further confusion. I am not convinced, however, that such terminological measures will prove to be very effective.

[7]  Of course, Callas's real name was 'Kalogeropoulos'.

purpose. Taken on its own, however, a name doesn't refer any more than a tie does.

It is instructive to compare names with number terms. Indeed, although number terms are rarely classified among the names, they are often used as such. Bearing a name may be likened to having a number: a crate in Rotterdam harbour may have a number stamped on it, for example 6. But of course the number term '6' doesn't refer to the crate in question, although it may be used to refer to it. That is to say, when taken on its own, a number term can sometimes be used to refer, but in most cases they must be part of larger expressions: *crate number 6*, *6 Elm Street*, *Paris VI*, *volume vi*, and so on. Names are similar to numbers, except that, in English at least, they can normally stand on their own when used to refer, and only sometimes must be embedded in a larger expression (as in *the river Thames*).

Another respect in which number terms are like names is that there countless ways in which an individual may be assigned a number or number term (the distinction is not always clear, but it doesn't seem to matter much, either). Sometimes individuals are inscribed with numbers: houses, banknotes, football players. In many other cases, there is a less immediate relation between a number and an individual that 'has it'. Numbers may be assigned at random or following a rigid procedure, for instance according to some linear order. There are various conventions for assigning house-numbers, convicts are numbered as they enter prison, cameras receive a number when they leave the factory, and in some countries citizens receive a number at birth.

Just as there many ways of assigning a number to an individual there are many ways of assigning it a name. There are descriptive names like *Fatty* or *Benjamin*. Similarly, the title of a book or film is expected to be somehow related to its content. In my native country, last names are assigned according to strict rules, while first names are afflicted by whim. Other cultures make use of patronymics, and in still others parents are named after a child. And so on. What all these naming practices have in common is just that some association is established between a name and its bearers, but how this association is initiated and sustained is different from case to case. The name-bearing relation between *Lolita* and the famous novel was initiated by the author and is sustained, inter alia, by printing the name on the front cover of every copy. The association between my last name and myself is sustained, inter alia, by records at a register office, but this does not apply for my first name. If I decided that I wanted to be called *Rudolf* instead of *Bart*, it would just be a nuisance to the people in my social sphere, but I don't think it would be humanly possible to change my last name into *Carnap* (not in my own country anyway).

It is not even true that the name-bearing relation is necessarily grounded in a social convention of some sort. For example, the files on my computer all

have names; some of these I have chosen myself, others have been chosen by various people around the globe, only some of whom I am acquainted with, and yet others, such as *cache183415.shtml* for example, were generated by some program. The association between a file and its name is entirely sustained by the computer's hardware and the programs running on it: if due to some programming error the file name *Bocherini* gets changed into *Corelli*, then the file's name IS *Corelli*, not *Bocherini*. The baptist's original intention is simply irrelevant.

The expression 'bearing a name' covers as many relations as there are naming practices, and it seems to me that Kripke's causal theory of reference is best viewed, not as a theory of reference, but as a partial theory of what it means to bear a name. The theory is only a partial one because there are naming practices that it doesn't account for, such as the one considered in the previous paragraph.

Kripke's charge that the quotation theory is circular can now be countered a second time, as follows. Even if a proponent of the quotation theory should want to explain how names can be used to refer, his analysis of names need not be circular, because the notion of reference need not enter into it: 'the individual named N' is not the same as 'the individual that N refers to', and the former does not presuppose the latter, either.

## 7.3 Names as presupposition inducers

On the presuppositional analysis of definiteness that I have taken more or less for granted in the foregoing chapters, a definite NP of the form 'the N' triggers the presupposition that N is given, and that is all there is to say about the meaning of definite NPs. Such NPs are purely presuppositional expressions, one might say, in that they presuppose their entire descriptive content (cf. § 1.4). On the presuppositional version of the quotation theory that I propose, names, too, are purely presuppositional expressions. The descriptive content of a name N is 'individual named N', and the name triggers the presupposition that an individual answering this description is given.

Let us see how this version of the quotation theory measures up to the observations presented in § 7.1. The first five of these observations hardly need any comment. Given that names are treated on a par with overt definite NPs, it doesn't come as a surprise that:

• names often carry definiteness markers;

• names are subject to definiteness restrictions;

• the referential/attributive distinction applies to names;

• names can be construed non-literally; and

• names can be construed generically.

I don't mean to suggest that I can account for definiteness restrictions, the distinction between referential and attributive readings, non-literal construals, and genericity. The point I want to make is merely that, given that definites are subject to definiteness restrictions and have referential, attributive, non-literal, and generic interpretations, the presuppositional theory of names leads us to expect that the same will hold for names, which it does.

The remaining observations from § 7.1 require more detailed explanations:

• Names are typically, though not always, used to refer to objects that are part of the common ground. That names prefer their referents to be given is a characteristic they have in common with all other presupposition inducers. We distinguished two ways of being part of the common ground: either the intended referent was already introduced into the discourse, or it is given in broader context in which the discourse takes place. To say that the presuppositions triggered by names are typically given in one of these ways is to say that they are rarely construed by means of accommodation. This is a peculiarity they share with many other presupposition inducing expressions, including some types of definite NPs; I will return to this point below.

• Names have bound-variable uses. *Any* presupposition that is linked to an antecedent in an embedded DRS will seem to behave like an element bound by a quantifier. So this is just a special case of the general rule that presuppositions prefer to be linked to a given antecedent.[8]

(14)    a.    If a child is christened 'Bambi', then Disney will sue Bambi's parents. (= (6b))

        b.    [: [x: child x, x is christened 'Bambi']
              $\Rightarrow$ [u: u is named 'Bambi', Disney sues u's parents]]

        c.    [: [x: child x, x is christened 'Bambi', x is named 'Bambi']
              $\Rightarrow$ [: Disney sues x's parents]]

The second occurrence of the name *Bambi* in (14a) triggers the presupposition that there is someone named *Bambi*, as shown in (14b) (at its first occurrence the name is mentioned, not used). This presupposition is bound in the antecedent of the conditional, and the resulting interpretation

---

[8]  Here I revert to the standard DRT analysis of conditionals, but only because it makes for DRSs that are easier to parse than the ones I have proposed (§ 6.4).

is (14c). Hence, on the present account there are are no relevant differences between (14a) and the following:

(15)  a.  If a farmer owns a donkey, he beats it.

  b.  If a farmer beats his donkey, his wife beats it, too.

In (15a) the pronoun refers back to the indefinite in the antecedent of the conditional, and similarly, in (15b) the presupposition triggered by the focus particle, that someone other than 'his' wife beats 'it', is bound in the antecedent of the conditional. The presuppositions triggered by *it* in the first sentence and *too* in the second are treated exactly the same as the presupposition induced by *Bambi* in (14a).

• Although names typically refer to entities in the common ground, they may be used to introduce new individuals, as in:

(16)  My best friend is John. (= (10))

In general, if a presupposition cannot be bound to a suitable antecedent, it will be accommodated. So unless there is a suitable antecedent for *John* in the context in which (16) is uttered (as when the sentence is preceded by, say, 'I have four good friends: John, Jack, Jim, and Sidney'), the presupposition that there is an individual with that name will be accommodated.

• Names generally take wide scope. It seems a rather safe conjecture that if a presupposition is bound it will usually be bound globally, i.e. in the principal DRS; for surely cases like (15a, b) are the exception rather than the rule. Moreover, if a presupposition must be accommodated, global accommodation is preferred to accommodation in an embedded DRS. Thus, on the whole presuppositions appear to have a strong tendency to be projected to the global level of the discourse representation. Since there is no reason to assume that this doesn't hold for names, we thus explain why names appear to prefer having wide scope. I say 'appear' because I reject the notion that definite NPs are scope-bearing expressions (see § 1.4).

The presuppositions triggered by names seem to have a decidedly stronger tendency to 'take wide scope' than some others. In this respect, too, they are on a par with other descriptively attenuate, 'incomplete', definites like *the door* or anaphoric pronouns like *it*. Compare the following sentences, for example:

(17)  a.  The President might have been a woman.

  b.  The car might have turned left.

(17a) allows for a reading according to which we might have had a situation *s* in which the President in *s* was a woman. In order to obtain this reading, the presupposition triggered by *the President* must be accommodated locally, i.e.

within the scope of the modal operator, and must not be linked to whoever happens to be president in the situation in which the sentence is uttered. This type of interpretation does not appear to be so readily available for (17b), presumably because *the car* insists more strongly that its referent be given (either in the previous discourse or in the wider context) than *the President* does. In this respect a proper name like *John* resembles *the car* more than *the President*, in that it dislikes being accommodated.

• Names can take narrow scope. Local accommodation is always dispreferred, and given that the presuppositions triggered by names already dislike being accommodated, it will be all the more difficult to think of a plausible scenario in which a name is interpreted by way of local accommodation. But, as we have seen in § 7.1, Bach managed to devise such a scenario:

(18)      Aaron Aardvark might have been president. (= (12))

Supposing that the hearer doesn't know of any person named *Aaron Aardvark*, and supposing that for whatever reason he doubts that the speaker knows of any such person, then the hearer will decide to locally accommodate the presupposition triggered by the name, which gives the correct reading.

And there is more evidence in favour of the presuppositional analysis of names. Suppose that shortly after John's birth his mother utters one of the following:

(20)      a.    If John had had red hair
          b.    If John had been a girl                his father would have
          c.    If John had been a twin                been even happier.
          d.    If John had been a Rolex watch

We should probably say that (20a) refers to a property John himself might have had. But being a Rolex watch is not such a property, and so (20d) is to be paraphrased rather along the following lines: 'If John's mother had given birth to a Rolex watch *instead of* to John (whatever that may mean), his father would have been even happier'. In short, while (20a) is about a counterfactual situation involving John, (20d) is about a counterfactual situation in which John's place has been taken by a Rolex watch. Note that, according to counterpart theory, which I still endorse, this distinction cannot be taken over into our semantic analysis, since no individual can inhabit more than one world. However, it does not follow that counterpart theory cannot account for the intuition, which I am trying to characterize, that there is such a distinction.

Examples (20b, c) are in a sense intermediate cases because it is less clear how we should describe them. Speaking only for myself: although I would probably say that John himself might have been a girl, I would rather not say

that he himself might have been a twin. Therefore, I would bracket (20b) with (20a) and (20c) with (20d); but I don't expect that these judgments will remain uncontested. This is as it may be, however, because the relevant observation is that our intuitions about the relation between John and his counterparts vary at all. Needless to say, this variation causes problems for a Kripkean analysis of names. In the present framework, by contrast, it is only to be expected.

The examples in (20) will all be interpreted along the following lines:

(21)   a.   [: [x: x is named 'John', ... ] ⇒ [x's father is happier]]

       b.   [: x is named 'John', [: ... ] ⇒ [x's father is happier]]

Let us assume for the nonce that the arrow symbolizes the counterfactual conditional; the antecedent of the conditional is filled in depending on which of (20a) through (20d) is under consideration. In (21a) the presupposition is triggered that there is an individual named *John*, and our presupposition theory predicts that this will project to the main DRS. The resulting interpretation is (21b), which says that there must a counterpart relation between John and the counterfactual individual satisfying the antecedent of the conditional. (21b) does not specify what kind of counterpart relation this should be, which is precisely what we want, because this relation varies from case to case, as we have seen in § 5.4. In other words, the variation we observed in (20) reflects the context dependence of the counterpart relation.

This analysis has two considerable merits. It gives a plausible explanation for the scale of readings exemplified by (20), and it explains how these interpretations come about without resorting to any special assumptions. The rigid-designator analysis, on the other hand, is just contradicted by these data.

### 7.4 The intuition of rigidity

In imitation of Bach (1987), I would have entitled this section 'The illusion of rigidity'. But on second thoughts I don't believe that rigidity is an illusion: it is a genuine empirical phenomenon. What is illusory is the notion that it calls for an explanation in semantic terms.

I have tried to explain why it is that names appear to have such a strong preference for taking wide scope. But Kripke's most forceful argument in favour of the thesis that names are rigid designators is based on the intuitive truth conditions of *simple* sentences, i.e. sentences without any relevant scope-bearing expressions (such as modals). The argument, which is only seemingly straightforward, as we will presently see, goes as follows. Consider a simple sentence containing a name, such as:

(22)      Mary is happy.

If someone understands this statement correctly, then he grasps a proposition which is true in a certain range of possible situations: in each of these situations, Mary is happy, where Mary is the referent of *Mary* in the situation in which (22) is uttered. This, Kripke suggests, *is* the claim that names are rigid designators.

  This observation appears to establish two points at the same time. First, since it is always the same person that makes the proposition expressed by (22) true, in any given possible situation, it would seem to prove that the name *Mary* is rigid. Secondly, it would seem to prove that the property of being named *Mary* isn't part of the meaning of the name, because the proposition expressed by (22) might be true even in a possible situation in which Mary had a different name. In brief, if Kripke's observation is correct, then names are rigid and the quotation theory is false.

  Kripke's observation is not as straightforward as it appears, however. To begin with, it matters a great deal how we frame our initial question. If we ask, as Kripke does, if (22) might be true in a possible situation in which Mary was called differently, then it will seem as if the name *Mary* is rigid. If, however, we ask what information is conveyed by this sentence, then it would seem to be part of its content that Mary is named *Mary*. The point I want to bring home is prefigured in Lewis's (1981) commentary on Kaplan's (1989) character/content distinction. Kaplan suggests that his notion of content captures the pre-theoretical notion of 'what is said' by uttering a sentence in a given context. For example, if two interlocutors simultaneously utter (23a) and (23b), respectively, what is said is the same:

(23)      a.    You are sleepy.

          b.    I am sleepy.

Or in Kaplan's terms: although these sentences have different characters, there are contexts in which their contents coincide.

  Lewis's objection is the following:

> Unless we give it some special technical meaning, the locution 'what is said' is very far from equivocal. [...] Kaplan's readers learn to focus on the sense of 'what is said' that he has in mind, ignoring the fact that the same words can be used to make different distinctions. For the time being, the words mark a definite distinction. But why mark that distinction rather than others that we could equally well attend to? (Lewis 1981: 97)

Lewis does not claim that Kaplan's theory is false. What he objects to is Kaplan's suggestion that his theoretical notion of content captures *the* pre-theoretical notion of what is said, and the implication that genuinely

alternative theories of meaning cannot capture this notion. This suggestion and its implication are wrong, because there is no single pre-theoretical notion of what is said. Intuitions about what is said vary with one's interests. Two interlocutors uttering (23a, b) may have said the same thing in one sense, but in another sense they haven't. A theory of lexical semantics may be expected to explain this shifty behaviour of the verb *say*, but there is no reason to require that central concepts in semantic theory (such as 'character' or 'content') must capture any of our ways of understanding 'what is said'.

Lewis's objection applies to Kripke's rigidity thesis, too. It is true that there is a sense in which (22) correctly describes a possible state of affairs in which Mary happily lives under a different name. But there is also a sense in which (22) is not correct in such a state of affairs; for if someone's name isn't *Mary*, it is *wrong* to call her so. Indeed, it has never been denied, as far as I am aware, that someone who utters (22) conveys the information that Mary is called *Mary*. What Kripke denies is merely that this information is part of the meaning of *Mary*, and he suggests that this equally holds for the pre-theoretical and theoretical notions of meaning. Given the chameleontic quality of the pre-theoretical notion of meaning, it is pointless to disagree with the first half of this claim; but the second half is false, in my view.

Still, why is it that names usually, though not invariably, appear to be rigid?[9] This is a legitimate question, even if we mustn't presuppose that the intuition of rigidity is a direct reflection of the semantics of proper names. My answer to this question is not a new one. It has been given, independently it seems, by several reliable sources, including Sommers (1982), Bach (1987), and Fodor (1987).[10] Of these, Fodor's version comes closest to the formulation that I prefer:

> The course of wisdom would be to reiterate the moral — viz., that names are a hard problem for everybody — and then to shut up and leave it alone. Still, how about this: 'Cicero' and 'Tully' are synonymous but differ in presupposition. [...] Then 'Cicero was wet' says, in effect, that *he was wet* and presupposes that he was called 'Cicero'. 'Tully was wet' says that *he was wet* too, but it presupposes that he was called 'Tully'. 'Cicero is Tully' is informative because, although it doesn't *say* that the guy who was called 'Cicero' was called 'Tully', it "carries the information" that he was. (Fodor 1987: 85)

---

[9] It should always be kept in mind that names aren't always rigid, as several of the data discussed in the foregoing have shown. This observation already suggests that the intuition of rigidity calls for a pragmatic explanation, not a semantic one.

[10] More accurately: it seems to me that the intuition underlying the answers given by these authors is the same in each case, although strictly speaking their proposals are incompatible with one another.

I would prefer putting it as follows. The meaning of a name N is 'the individual named N', where the semantic contribution of the definite article is to be analysed in presuppositional terms. Given that the property of bearing the name N is an accidental one (I could name my left ear *Fortinbras* if I chose), referring to an individual with N will not, in general, be particularly effective unless the hearer already knows the intended referent and that it is named N. Therefore, a name will practically always be used to refer to an individual that was already given to the hearer beforehand, and if a name is thus used it will appear to be rigid.

To say that 'the N', where N stands for any nominal head, is a presuppositional expression is to say that its intended referent *a* is presented as given, and that it is presented as given that *a* is an N. Someone who utters (22), for example, presupposes that there is a person *a* whose name is *Mary* and asserts that *a* is happy. So what Kripke calls the 'meaning' of this sentence is just its asserted content, and it is of course correct that if we confine our attention to this part of the information  conveyed by (22), the 'meaning' of *Mary* will appear to be rigid, and the property of 'being named 'Mary' ' will appear to be irrelevant.

As Bach (1987) points out, the strategy which Kripke consistently employs in *Naming and Necessity* is to first fix the referent of a name before he broaches the issue whether or not the name is rigid. But then it shouldn't come as a surprise that names always turn out to be rigid. Nor is it surprising that, for Kripke, non-referring names *eo ipso* disqualify as admissible evidence.

My suggestion is that names appear to be rigid because they are presuppositional expressions. But surely not all presuppositional expressions appear to be rigid designators. So there must be something else that gives names their peculiar rigid feel. There is, but it is not peculiar to names. Consider the following situation. John and I are standing in the corridor, discussing the weather. We have just witnessed Mary walk past us and enter the kitchen. A moment later we hear cheerful singing emanating from the kitchen. At that point I say:

(24)     She is happy.

To John it will seem as if *she* is a rigid designator. For what I have said is true iff Mary is happy, and this is what it means to say, according to Kripke, that an expression is rigid (cf. Sommers 1982). Why does it seem to John as if *she* is rigid? According to DRT, John maintains a representation of the speaker's commitment slate, which upon my utterance of (24) might look as follows (I ignore the previous discourse because it was irrelevant to my utterance of (24)):

(25)     [x: person x, female x, happy x]

Since John treats anaphora as a special case of presupposition, he assumes that the pronoun triggers the presupposition that there is a given person who is female, and that it is asserted that this person is happy. There is no suitable antecedent for this presupposition to be bound to, so it must be accommodated. But even so John will try to integrate the presupposition with his other knowledge, which is to say, in the present case, that he will connect it to his representation of Mary. Thus John's belief box, in which his representation of my commitment slate is embedded, may be pictured as follows:

(26)    [x: woman x, x is named 'Mary', x is in the kitchen,
        Bart believes: [x: person x, female x, happy x]]

(This representation abstracts away from the concerns that engaged us in Chapter 5.) John believes that there is a woman called *Mary*, that she is in the kitchen, and that I believe of this person that she is a female person who is happy. Or rather, the person that John's beliefs are about and the person that he takes my statement to be about are counterparts. This is why the pronoun in (24) appears to be rigid.

By uttering (24) I have conveyed the information in (25), part of which is that Mary is a female person; for this is the descriptive content of the pronoun *she* (by approximation, at least). But when we ask John if what I have said might be true even if it transpired that Mary is a man, his answer will be yes. This is a reasonable answer because one very common way of interpreting 'what is said' is by restricting it to asserted information.

If instead of (24) I had uttered (22), John would construe the name as 'rigid' for the same reason he would construe the pronoun in (24) as 'rigid'. And he would say that the property of being named *Mary* is not part of what I have said for the same reason he would say that the property of being a female person is not part of the statement made by (24).

### 7.5  Concluding remarks

It is always possible to misconstrue a story, no matter how simple, and I know from experience that my story about names is especially liable to be misunderstood, so let me briefly mention some of the things I didn't claim and don't want to claim, either.

First, I don't want to claim that names are pronouns. This is how Sommers (1982) puts it, but not I. I do want to claim that names are very much like pronouns. They both are presuppositional expressions, and have the full complement of possible interpretations that presuppositional expressions generally allow for. But, of course, the descriptive content of a pronoun is of a different category than that of a name, which makes it much easier to construct examples of 'blocking' for pronouns than for names. For example:

(27)      If Mary has a car, it is pink.

It doesn't require much ingenuity to devise examples like (27). Analogous examples with names are much more difficult to construct. The essential characteristic that names and pronouns share is that they are nearly always used to refer to an object that speaker and hearer take to be given (with names this tendency is even stronger than with pronouns). It is this pragmatic fact which accounts for the intuition of rigidity.

Secondly, I didn't claim that names and pronouns appear to be rigid because they generally link up to antecedents that are. I might have claimed this, because the main point that I wanted to establish in this chapter is that the quotation theory is right, and this is a theory about the meaning of names. But I didn't.

Thirdly, I didn't claim that names and pronouns are the only types of expressions that usually seem to be rigid. For example, there are many overt definites that behave exactly the same way, the most obvious class being semantically attenuate descriptions like *the man*, *the thing*, and so on. Also, I believe that the presuppositions associated with quantifiers like *all* and *most* are often rigid in the same sense in which names are.

(28)      Everybody is happy.

Typically, this will be uttered in a situation in which the domain of *everybody* is contextually given. Let c be a context in which (28) is uttered, and let A be the intended domain of *everybody* in c. The proposition expressed by (28) in c is true in any given situation c' iff all individuals in A are happy in c'. Hence, *everybody* is rigid — or, better perhaps, its interpretation in c has a rigid component.

On the other hand, I don't want to claim that all presuppositional expressions engender the impression of rigidity. To the extent that we have clear intuitions about rigidity at all they pertain to individuals and sets of individuals. Moreover, the intuition that an expression $\alpha$ is rigid is stronger when we feel that the descriptive content of $\alpha$ is inessential. It is for these reasons that we wouldn't want to say that factives are rigid, although *qua* presuppositional expressions they are quite similar to definite NPs.

At the outset I announced that I would try to rehabilitate Kneale's analysis of proper names. But didn't I actually replace it with something completely different? I don't think so. Kneale's crucial insight was that names are synonymous with definite NPs of the form 'the individual named so and so'. Of course, this is not yet a theory of names; but it has the merit of broadening the problem. If at this point Kneale had adopted Russell's theory of descriptions, his theory of names would already have been superior to the rigid-designator account, as I have tried to show in the first half of this chapter. But Kneale actually suggested a *presuppositional* analysis of definite

NPs, and in doing so he broadened the problem even further, although this will not have been evident at the time.

What I have done is implement Kneale's idea in a theory of presupposition that is more explicit and, presumably, more comprehensive than what he envisaged. The resulting analysis of names is superior to Kripke's for two main reasons. First, it explains a whole range of empirical facts about names that Kripke's theory cannot account for, if they don't falsify it to begin with. Secondly, it treats names as just another class of presuppositional expressions within the framework of a theory that is amply motivated on independent grounds. In comparison, Kripke's proposal is definitely *ad hoc*.

*This page intentionally left blank*

# References

Asher, N. 1986: Belief in discourse representation theory. *Journal of Philosophical Logic* 15: 127–189.

Asher, N. 1987: A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10: 125–197.

Asher, N. 1993: *Reference to Abstract Objects in Discourse*. Kluwer, Dordrecht.

Atlas, J.D. and S. Levinson 1981: *It*-clefts, informativeness and logical form: radical pragmatics (revised standard version). In: P. Cole (ed.), *Radical Pragmatics*. Academic Press, New York. Pp. 1–61.

Bach, K. 1981: What's in a name. *Australasian Journal of Philosophy* 59: 371–386.

Bach, K. 1987: *Thought and Reference*. Oxford University Press.

Bakker, R. 1988: *The Dinosaur Heresies*. Penguin Books, Harmondsworth. First edition, 1986.

Barwise, J. 1987: Noun phrases, generalized quantifiers, and anaphora. In: P. Gärdenfors (ed.), *Generalized Quantifiers: Linguistic and Logical Approaches*. Reidel, Dordrecht. Pp. 1–29.

Beaver, D.I. 1992: The kinematics of presupposition. *Proceedings of the 8th Amsterdam Colloquium*. University of Amsterdam. Pp. 17–36.

Beaver, D.I. 1993: *What Comes First in Dynamic Semantics*. ILLC Prepublication Series LP-93–15, University of Amsterdam.

Beaver, D.I. 1994: When variables don't vary enough. In: M. Harvey and L. Santelmann (eds.), *Proceedings from Semantics and Linguistic Theory IV*. Cornell University, Ithaca. Pp. 35–60.

Beaver, D.I. 1997: Presupposition. In: J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Linguistics*. Elsevier, Amsterdam, and MIT Press, Cambridge, Mass. Pp. 939–1008.

Belnap, N. 1970: Conditional assertion and restricted quantification. *Noûs* 4: 1–12.

Boër, S.E. and W.G. Lycan 1976: *The Myth of Semantic Presupposition*. Indiana University Linguistics Club, Bloomington.

Bolinger, D. 1977: *Pronouns and Repeated Nouns*. Indiana University Linguistics Club, Bloomington.

Bosch, P. 1983: *Agreement and Anaphora*. Academic Press, London.

Bosch, P. 1988: Representing and accessing focussed referents. *Language and Cognitive Processes* 3: 207–231.

Bosch, P. and B. Geurts 1990: Processing definite NPs. *Rivista di Linguistica* 2: 177–199.

Carston, R. 1996: Metalinguistic negation and echoic use. *Journal of Pragmatics* 25: 309–330.

Chierchia, G. 1994: Intensionality and context change: towards a dynamic theory of propositions and properties. *Journal of Logic, Language, and Information* 3: 141–168.

Chierchia, G. 1995: *Dynamics of Meaning*. University of Chicago Press.

Chierchia, G. and S. McConnell-Ginet 1990: *Meaning and Grammar*. MIT Press, Cambridge, Mass.

Comrie, B. 1986: Conditionals: a typology. In: Traugott et al. (eds.), 77–99.

Cresswell, M. 1985: *Structured Meanings: The Semantics of Propositional Attitudes*. MIT Press, Cambridge, Mass.

Davis, S. (ed.) 1991: *Pragmatics: A Reader*. Oxford University Press.

Dekker, P. 1993: *Transsentential Meditations*. Doctoral dissertation, University of Amsterdam.

Delacruz, E.B. 1976: Factives and propositional level constructions in Montague grammar. In: B.H. Partee (ed.), *Montague Grammar*. Academic Press, New York.

Douloureux, P.R.N.T. 1971: A note on one's privates. In: A.M. Zwicky, P.H. Salus, R.I. Binnick, and A.L. Vanek (eds.), *Studies out in Left Field: Defamatory Essays Presented to James D. McCawley on the Occasion of his 33rd or 34th Birthday*. Benjamins, Amsterdam (1992 reprint). Pp. 45–51.

Edelberg, W. 1992: Intentional identity and the attitudes. *Linguistics and Philosophy* 15: 561–596.

Evans, G. 1977: Pronouns, quantifiers, and relative clauses (I). *Canadian Journal of Philosophy* 7: 467–536.

Fauconnier, G. 1985: *Mental Spaces*. MIT Press, Cambridge, Mass.

Fodor, J.A. 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press, Cambridge, Mass.

Fodor, J.D. and I.A. Sag 1982: Referential and quantificational indefinites. *Linguistics and Philosophy* 5: 355–398.

Frank, A. 1997: *Context Dependence in Modal Constructions*. Doctoral dissertation, University of Stuttgart.

Gazdar, G. 1979: *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press, New York.

Geach, P.T. 1962: *Reference and Generality*. Cornell University Press, Ithaca.

Geach, P.T. 1967: Intentional identity. *Journal of Philosophy* 64: 627–632.

Geurts, B. 1985: Generics. *Journal of Semantics* 4: 257–270.

Geurts, B. 1988: The representation of generic knowledge. In: M. Krifka (ed.), *Genericity in Natural Language*. SNS-Bericht 88–42, Seminar für natürlich-sprachliche Systeme, University of Tübingen. Pp. 155–174.

Geurts, B. 1995: *Presupposing*. Doctoral dissertation, University of Stuttgart.

Geurts, B. 1996a: Local satisfaction guaranteed: a presupposition theory and its problems. *Linguistics and Philosophy* 19: 259–294.

Geurts, B. 1996b: On 'no'. *Journal of Semantics* 13: 67–86.

Geurts, B. 1996c: A presuppositional account of domain restriction. In: D. Gibbon (ed.), *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference*. Mouton de Gruyter, Berlin. Pp. 293–304.

Geurts, B. 1997a: Dynamic Dido and commutative Aeneas. In: R. van der Sandt, R. Blutner, and M. Bierwisch (eds.), *From Underspecification to Interpretation: Papers from the ASG Workshop*. ILL Working Paper 29, Institute for Logic and Linguistics, IBM Scientific Centre, Heidelberg. Pp. 163–182.

Geurts, B. 1997b: Dynamic semantics vs. DRT. *Zeitschrift für Sprachwissenschaft* 16: 209–226.

Geurts, B. 1997c: Good news about the description theory of names. *Journal of Semantics* 14: 319–348.

Geurts, B. 1998a: The mechanisms of denial. *Language* 74: 274–307.

Geurts, B. 1998b: Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21: 545–601.

Geurts, B. and R. van der Sandt 1997: Presuppositions and backgrounds. *Proceedings of the 11th Amsterdam Colloquium*. University of Amsterdam. Pp. 37–42.

Geurts, B. and R. van der Sandt, 1999: Domain restriction. In: P. Bosch and R. van der Sandt (eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge University Press. Pp. 268–292.

Greenberg, J.H. 1963: Some universals of grammar with particular reference to the order of meaningful elements. In: J.H. Greenberg (ed.), *Universals of Language*. MIT Press, Cambridge, Mass. Pp. 73–113.

Grice, H.P. 1975: Logic and conversation. In: P. Cole and J.L. Morgan (eds.), *Speech Acts*. Academic Press, New York. Pp. 41–58. Reprinted in: Grice (1989), 22–40.

Grice, H.P. 1981: Presupposition and conversational implicature. In: P. Cole (ed.), *Radical Pragmatics*. Academic Press, New York. Pp. 183–198. Reprinted in: Grice (1989), 269–282.

Grice, H.P. 1989: *Studies in the Way of Words*. Harvard University Press, Cambridge, Mass.

Grimm, R. and R. Merrill (eds.) 1987: *Contents of Thought*. University of Arizona, Tucson.

Groenendijk, J. and M. Stokhof 1989: Context and information in dynamic semantics. In: H. Bouma and B.A.G. Elsendoorn (eds.), *Working Models of Human Perception*. Academic Press, London. Pp. 457–486.

Groenendijk, J. and M. Stokhof 1991: Dynamic predicate logic. *Linguistics and Philosophy* 14: 39–100.

Groenendijk, J. and M. Stokhof 1997: Questions. In: J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Linguistics*. Elsevier, Amsterdam, and MIT Press, Cambridge, Mass. Pp. 1055–1124.

Gundel, J.K., N. Hedberg, and R. Zacharski 1993: Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307.

Haiman, J. 1978: Conditionals are topics. *Language* 54: 564–589.

Haiman, J. 1986: Constraints on the form and meaning of the protasis. In: Traugott et al. (eds.), 215–227.

Hamblin, C.L. 1971: Mathematical models of dialogue. *Theoria* 37: 130–155.

Hamblin, C.L. 1987: *Imperatives*. Blackwell, Oxford.

Harper, W.L., R. Stalnaker, and G. Pearce (eds.) 1981: *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Reidel, Dordrecht.

Haviland, S. and H.H. Clark 1974: What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 13: 512–521.

Heim, I. 1982: *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst.

Heim, I. 1983: On the projection problem for presuppositions. *Proceedings of the West Coast Conference on Formal Linguistics* 2: 114–126. Reprinted in: Davis (ed.) 1991: 397–405.

Heim, I. 1990: E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13: 137–178.

Heim, I. 1992: Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9: 183–221.

Hintikka, J. 1962: *Knowledge and Belief*. Cornell University Press, Ithaca, N.Y.

Hintikka, J. 1969: Semantics for propositional attitudes. In: J.W. Davis, D.J. Hockney, and W.K. Wilson (eds.), *Philosophical Logic*. Reidel, Dordrecht. Pp. 21–45.

Hintikka, J. 1983: New foundations for a theory of questions and answers. In: F. Kiefer (ed.), *Questions and Answers*. Reidel, Dordrecht. Pp. 159–190.

Hirst, G. 1981: *Anaphora in Natural Language Understanding: A Survey*. Springer-Verlag, Berlin.

Horn, L.R. 1989: *A Natural History of Negation*. Chicago University Press.

Kadmon, N. 1990: Uniqueness. *Linguistics and Philosophy* 13: 273–324.

Kamp, H. 1981: A theory of truth and semantic representation. In: J.A.G. Groenendijk, T.M.V. Janssen, and M.B.J. Stokhof (eds.), *Formal Methods in the Study of Language*. Mathematical Centre Tracts 135, Amsterdam. Pp. 277–322.

Kamp, H. 1987: Comments [on Stalnaker 1987]. In: Grimm and Merrill (eds.), 156–181.

Kamp, H. 1988: Conditionals in DR theory. In: J.P. Hoepelman (ed.), *Representation and Reasoning*. Niemeyer, Tübingen. Pp. 66–124.

Kamp, H. 1990: Prolegomena to a structural theory of belief and other attitudes. In: C.A. Anderson and J. Owens (eds.), *Propositional Attitudes: The Role of Content in Language, Logic, and Mind*. CSLI, Stanford. Pp. 27–90.

Kamp, H. and U. Reyle 1993: *From Discourse to Logic*. Kluwer, Dordrecht.

Kaplan, D. 1969: Quantifying in. In: D. Davidson and J. Hintikka (eds.), *Words and Objections*. Reidel, Dordrecht. Pp. 206–242.

Kaplan, D. 1989: Demonstratives: an essay on the semantics, logic, metaphysics, and episte- mology of demonstratives and other indexicals. In: J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. Oxford University Press. Pp. 481–563.

Karttunen, L. 1969: Pronouns and variables. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, 108–115.

Karttunen, L. 1973a: Presuppositions of compound sentences. *Linguistic Inquiry* 4: 167–193.

Karttunen, L. 1973b: The last word. Mimeograph, University of Texas, Austin. Reference as given by Heim (1992).

Karttunen, L. 1974: Presupposition and linguistic context. *Theoretical Linguistics* 1: 181–194.

Karttunen, L. 1976: Discourse referents. In: J. McCawley (ed.), *Notes from the Linguistic Underground*. Academic Press, New York. Pp. 363–385.

Karttunen, L. and S. Peters, 1979: Conventional implicature. In: C.-K. Oh and D. Dinneen (eds.), *Presupposition*. Academic Press, New York. Pp. 1–56.

Kay, P. 1992: The inheritance of presuppositions. *Linguistics and Philosophy* 15: 333–379.

Keenan, E.L. and R.D. Hull 1973: The logical presuppositions of questions and answers. In: J.S. Petöfi and D. Franck (eds.), *Presuppositions in Philosophy and Linguistics*. Athenäum Verlag, Frankfurt. Pp. 441–466.

Kempson, R.M, 1977: *Semantic theory*. Cambridge University Press.

Kempson, R.M. 1975: *Presupposition and the Delimitation of Semantics*. Cambridge University Press.

Kneale, W. 1962: Modality de dicto and de re. In: E. Nagel, P. Suppes, and A. Tarski (eds.), *Logic, Methodology and Philosophy of Science. Proceedings of the 1960 International Congress*. Stanford University Press. Pp. 622–633.

Krahmer, E. 1993: Partiality and dynamics: theory and application. *Proceedings of the 9th Amsterdam Colloquium*. University of Amsterdam. Pp. 391–410.

Krahmer, E. and R. Muskens 1995: Negation and disjunction in discourse representation theo- ry. *Journal of Semantics* 12: 357–376.

Kratzer, A. 1979: Conditional necessity and possibility. In: R. Bäuerle, U. Egli, and A. von Stechow (eds.), *Semantics from Different Points of View*. Springer-Verlag, Berlin.

Kratzer, A. 1981: The notional category of modality. In: H. Eikmeyer and H. Rieser (eds.), *Words, Worlds, and Contexts*. De Gruyter, Berlin. Pp. 38–74.

Kratzer, A. 1991: Modality. In: A. von Stechow and D. Wunderlich (eds.), *Semantics: An International Handbook of Contemporary Research*. De Gruyter, Berlin. Pp. 639–650.

Krifka, M. 1993: Focus and presupposition in dynamic interpretation. *Journal of Semantics* 10: 269–300.

Kripke, S.A. 1977: Speaker's reference and semantic reference. In: P.A. French, T.E. Uehling, and H.K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*. University of Minnesota Press, Minneapolis.

Kripke, S.A. 1980: *Naming and Necessity*. Blackwell, Oxford. First published in: D. Davidson and G. Harman (eds.) 1972: *Semantics for Natural Language*. Reidel, Dordrecht.

Kuroda, S.-Y. 1979: Katz and Langendoen on presupposition. In: C.-K. Oh and D. Dinneen (eds.), *Presupposition*. Academic Press, New York. Pp. 183–198.

Lakoff, G. 1972: Linguistics and natural logic. In: D. Davidson and G. Harman (eds.), *Semantics of Natural Language*. Reidel, Dordrecht. Pp. 545–665.

Lakoff, G. and J.R. Ross 1972: A note on anaphoric islands and causatives. *Linguistic Inquiry* 3: 121–125.

Lamarque, P.V. 1997: *Concise Encyclopedia of Philosophy of Language*. Elsevier, Oxford.

Levinson, S.C. 1983: *Pragmatics*. Cambridge University Press.

Levinson, S.C. 1987: Pragmatics and the grammar of anaphora: a partial pragmatic reduction of Binding and Control phenomena. *Journal of Linguistics* 23: 379–434.

Levinson, S.C. 1997: *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Ms., Max Planck Institute for Psycholinguistics, Nijmegen.

Lewis, D.K. 1968: Counterpart theory and quantified modal logic. *Journal of Philosophy* 65: 113–126. Reprinted in: Lewis (1983a), 26–39.

Lewis, D.K. 1971: Counterparts of persons and their bodies. *Journal of Philosophy* 68: 203–211. Reprinted in: Lewis (1983a), 47–54.

Lewis, D.K. 1973: *Counterfactuals*. Harvard University Press, Cambridge, Mass.

Lewis, D.K. 1975: Adverbs of quantification. In: E.L. Keenan (ed.), *Formal Semantics of Natural Language*. Cambridge University Press. Pp. 3–15.

Lewis, D.K. 1979: Scorekeeping in a language game. *Journal of Philosophical Logic* 8: 339–59. Reprinted in: Davis (ed.) 1991: 416–427.

Lewis, D.K. 1980: Index, context, and content. In: S. Kanger and S. Öhman (eds.), *Philosophy and Grammar*. Reidel, Dordrecht. Pp. 79–100.

Lewis, D.K. 1983a: *Philosophical Papers*. Volume 1. Oxford University Press.

Lewis, D.K. 1983b: Postscripts to 'Counterpart theory and quantified modal logic'. In: Lewis (1983a), 39–46.

Liberman, M. 1973: Alternatives. *Papers from the 9th Regional Meeting of the Chicago Linguistic Society*, 416–27.

Loar, B. 1972: Reference and propositional attitudes. *Philosophical Review* 81: 43–62.

Loar, B. 1976: The semantics of singular terms. *Philosophical Studies* 30: 353–377.

Lycan, W.G. 1984: *Logical Form in Natural Language*. MIT Press, Cambridge, Mass.

Lycan, W.G. 1987: The myth of the 'projection problem for presupposition'. *Philosophical Topics* 15: 169–175.

McCawley, J.D. 1981: *Everything that Linguists have Always Wanted to Know about Logic but were Ashamed to Ask*. Blackwell, Oxford.

Mercer, R.E. 1992: Default logic: towards a common logical semantics for presuppositions and entailments. *Journal of Semantics* 9: 223–250.

Montague, R. 1973: The proper treatment of quantification in ordinary English. In: K.J.J. Hintikka, J.M.E. Moravcsik, and P. Suppes (eds.), *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*. Reidel, Dordrecht. Pp. 221–242.

Morgan, J. 1969: On the treatment of presupposition in transformational grammar. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, 167–177.

Muskens, R. 1990: Anaphora and the logic of change. In: J. van Eijck (ed.), *JELIA '90, European Workshop on Logics in AI*. Springer, Berlin. Pp. 414–430.

Neale, S. 1990: *Descriptions*. MIT Press, Cambridge, Mass.

Neale, S. 1992: Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15: 509–559.

Quine, W.V. 1950: *Methods of Logic*. Holt, Rinehart, and Winston, New York.

Quine, W.V. 1956: Quantifiers and propositional attitudes. *Journal of Philosophy* 53: 177–187.

Ramsey, F.P. 1931: General propositions and causality. In: *The Foundations of Mathematics and Other Logical Essays*. Routledge, London. Pp. 237–255.

Reyle, U. 1993: Dealing with ambiguities by underspecification: construction, representation and deduction. *Journal of Semantics* 10: 123–179.

Roberts, C. 1987: *Modal Subordination, Anaphora and Distributivity*. Ph.D. thesis, University of Massachusetts, Amherst.

Roberts, C. 1989: Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12: 683–721.

Russell, B. 1905: On denoting. *Mind* 14: 479–493.

Soames, S. 1979: A projection problem for speaker presuppositions. *Linguistic Inquiry* 10: 623–666.

Soames, S. 1982: How presuppositions are inherited: a solution to the projection problem. *Linguistic Inquiry* 13: 483–545. Reprinted in: Davis (ed.) 1991: 428–470.

Soames, S. 1989: Presupposition. In: D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, Volume 4*. Reidel, Dordrecht. Pp. 553–616.

Sommers, F. 1982: *The Logic of Natural Language*. Clarendon Press, Oxford.

Stalnaker, R.C. 1968: A theory of conditionals. In: N. Rescher (ed.), *Studies in Logical Theory*. Blackwell, Oxford. Pp. 98–112. Reprinted in: Harper, Stalnaker, and Pearce (eds.) 1981: 41–55.

Stalnaker, R.C. 1970: Pragmatics. *Synthese* 22: 272–289.

Stalnaker, R.C. 1973: Presuppositions. *Journal of Philosophical Logic* 2: 447–457.

Stalnaker, R.C. 1974: Pragmatic presuppositions. In: M.K. Munitz and P.K. Unger (eds.), *Semantics and Philosophy*. New York University Press. Pp. 197–213. Reprinted in: Davis (ed.) 1991: 471–482.

Stalnaker, R.C. 1975: Indicative conditionals. *Philosophia* 5: 269–286. Reprinted in: Harper, Stalnaker, and Pearce (eds.) 1981: 193–210.

Stalnaker, R.C. 1978: Assertion. In: P. Cole (ed.), *Pragmatics*. Academic Press, New York. Reprinted in: Davis (ed.) 1991: 278–289.

Stalnaker, R.C. 1981: A defense of conditional excluded middle. In: Harper, et al. (eds.), 87–104.

Stalnaker, R.C. 1984: *Inquiry*. MIT Press, Cambridge, Mass.

Stalnaker, R.C. 1987: Belief attribution and context. In: Grimm and Merrill (eds.), 141–156.

Strawson, P.F. 1950: On referring. *Mind* 59: 320–344.

Strawson, P.F. 1964: Identifying reference and truth-values. *Theoria* 30: 96–118.

Traugott, E.C., A. ter Meulen, J.S. Reilly, and C.A. Ferguson (eds.) 1986: *On Conditionals*. Cambridge University Press.

Turner, K. 1992: Defending semantic presupposition. *Journal of Pragmatics* 18: 345–371.

Van Deemter, K. 1990: Forward references in natural language. *Journal of Semantics* 7: 281–300.

Van der Sandt, R.A. 1987: Presupposition and discourse structure. *Proceedings of the 6th Amsterdam Colloquium*. University of Amsterdam. Pp. 312–330. Reprinted with revisions in: R. Bartsch, J. van Benthem, and P. van Emde Boas (eds.) 1989: *Semantics and Contextual Expresssion*. Foris, Dordrecht. Pp. 267–294.

Van der Sandt, R.A. 1988: *Context and Presupposition*. Routledge, London.

Van der Sandt, R.A. 1991: Denial. *Papers from the 27th Regional Meeting of the Chicago Linguistic Society, Part 2: The Parasession on Negation*, 331–344.

Van der Sandt, R.A. 1992: Presupposition projection as anaphora resolution. *Journal of Semantics* 9: 333–377.

Van der Sandt, R.A. and B. Geurts 1991: Presupposition, anaphora, and lexical content. In: O. Herzog and C.-R. Rollinger (eds.), *Text Understanding in LILOG*. Springer-Verlag, Berlin. Pp. 259–296.

Van Eijck, J. 1993: The dynamics of description. *Journal of Semantics* 10: 239–267.

Van Fraassen, B. 1969: Presuppositions, super-valuations, and free logic. In: K. Lambert (ed.), *The Logical Way of Doing Things*. Yale University Press, New Haven. Pp. 67–91.

Veltman, F. 1996: Defaults in update semantics. *Journal of Philosophical Logic* 25: 221–261.

Walker, R.C.S. 1975: Conversational implicatures. In: S. Blackburn (ed.), *Meaning, Reference and Necessity*. Cambridge University Press. Pp. 133–181.

Wilson, D. 1975: *Presuppositions and Non-Truth-Conditional Semantics*. Academic Press, London.

Zeevat, H. 1984: Belief. In: F. Landman and F. Veltman (eds.), *Varieties of Formal Semantics*. Foris, Dordrecht. Pp. 405–425.

Zeevat, H. 1989a: A compositional approach to discourse representation theory. *Linguistics and Philosophy* 12: 95–131.

Zeevat, H. 1989b: Realism and definiteness. In: G. Chierchia, B.H. Partee, and R. Turner (eds.), *Properties, Types and Meaning, Vol. 2*. Kluwer, Dordrecht. Pp. 269–297.

Zeevat, H. 1991: *Aspects of Discourse Semantics and Unification Grammar*. Doctoral dissertation, University of Amsterdam.

Zeevat, H. 1992: Presupposition and accommodation in update semantics. *Journal of Semantics* 9: 379–412.

Zeevat, H. 1996: A neoclassical analysis of belief sentences. *Proceedings of the 10th Amsterdam Colloquium*. University of Amsterdam. Pp. 723–742.

*This page intentionally left blank*

# Subject index

# Index of names