IDs IN SYNTAX AND DISCOURSE:
An Analysis of Extraposition in German

DISSERTATION

zur Erlangung des Doktorgrades
an der
geisteswissenschaftlichen Fakultät
der Universität Wien

eingereicht von
Martina WILTSCHKO

Wien, Juni 1995

Acknowledgements

It is a pleasure to have Edwin Williams as a supervisor. I want to thank him for giving me the chance to spend a semester in Princeton and for accepting to supervise this thesis. It is a great and important experience to study with him - his way of looking at things was of great influence to me. And also: it was fun!

I am grateful to W.U.Dressler not only for being on my committee but also for providing a very fruitful and stimulating atmosphere at the department of linguistics in Vienna. His critical attitude towards the generative theory has always been a challenge - moreover it gave me the possibility to look at linguistics from a variety of perspectives.

Vienna is a great place to study linguistics! Many thanks to Martin Prinzhorn for this situation. Before his "Einführung in die spezielle Grammatiktheorie" in 1989 I did not believe that Grammar is an interesting area - since then I completely changed my mind and Martin is responsible for this to an immense degree. He is not only a great teacher, but also a great friend. This combination makes it possible that he always seems to know what is best for his students. In my case he knew that I should visit Princeton for a semester - and he was right!

The department in Vienna would not be what it is without my friends and colleagues: Elena Anagnostopoulou, Gerhard Brugger, Christine Czinglar, Fuzi Fultin, Martin Hackl, Martin Haiden, Katharina Köhler, Robert Kühnel, Winfried Lechner, Friedrich Neubarth, Birgit Trinker, Susi Wurmbrand. Thanks also to my colleagues and (almost) office-mates: Chris Schaner, Jacek Witkos, Peter Bodor, Mario Palaschke and Sigrid Adam.

Special thanks to Strang Burton not only for fruitful discussions and various comments, questions and "you-have-to-completely-rewrite-this"-remarks but also for his friendship, moral support and listening to (and reading) my complaints about the impossibility to ever write a dissertation. (Answers might have taken a little longer across the Atlantic...)

Finally, thanks to my family. Without their support I would have never been able to study linguistics in the first place.

TABLE OF CONTENTS

INTRODUCTION

The main goal of this dissertation is to give an empirical and theoretical adequate account of a construction which is traditionally referred to as *extraposition*. Any fruitful analysis within the generative framework has in general been an attempt to provide an answer to how a particular element and/or relation respectively is licensed. The main question to be addressed here is therefore: how is the extraposed constituent licensed? I will try to give an account within the Principles & Parameters framework (in the sense of Chomsky 1981), which I assume the reader to be familiar with.

The main challenge of such an approach is to get rid of construction specific statements. This is a necessary consequence of the assumption that there are no such things as particular 'constructions'. Terms like 'extraposition', 'passive', 'relative clauses', etc. are considered to be more or less abbreviations capturing some empirical phenomena. It seems to me that most of the traditional analyses for extraposition cannot overcome some construction-specific statements. The most striking fact that concerns every single analysis is the stipulative character of *rightness*. Rightness within syntactic (hierarchical) configurations however should not play any role at all, since for hierarchical configuration left- or rightness does not change anything. The main goal of the analysis presented here is to get rid of the notion of rightness w.r.t. extraposition. We will see how rightness once replaced by the notion of *precedence* (which is a crucial notion for linear ordering rather than hierarchical ordering) can be derived by some version of the *Novelty Condition* in the sense of Heim (1982).

Many problems seem to arise from this crucial point of my analysis. I will however show that these apparent problems can be solved. I will show that the assumptions that I will make in order to solve these apparent problems are independently motivated. In solving the apparent problems many other phenomena (which are not related to extraposition in the first place) will be addressed. Some of those phenomena will be treated more carefully, others will just be sketched. I will introduce the notion of *Identification* as a primitive (licensing-) relation which plays a crucial role in the theory of Grammar, not only for licensing extraposed constituents but also for other non-related phenomena. It is not unexpected that in treating one particular construction and postulating a certain kind of licensing relation for this construction we find other (unrelated) constructions which show the same sort of licensing relation (although instantiated differently due to the interaction of independent principles and parameters). This will however turn out to be a nice result since within a Principles & Parameters framework constructions itself are the mere output of such

interacting principles. Thus if we would just postulate a certain kind of licensing requirement for a certain kind of construction we would be back to a construction specific statement. However, it is not the purpose of this thesis to give a full-fledged analysis of Identification in general, this would go far beyond the scope of this work. Therefore sometimes I will just sketch some possible implications of the present analysis for some other constructions, without going into detail or working out the predictions.

This thesis is structured as follows. In the first chapter I will address issues concerning the theory of licensing. I will argue that Identification is a licensing relation. In chapter I. I will show how Identification as a licensing relation fits into the general theory of licensing. I will discuss the nature of and also the restrictions on various licensing relations. Moreover I will show that the notion of Identification is not a new concept. Other phenomena are analyzed as involving this very licensing requirement. Once those general ideas are discussed I will turn to some particular problems concerning licensing w.r.t. extraposition - thus we will turn to some particular problems w.r.t. extraposition and previous analyses for this construction. This will then be the starting point to present the analysis suggested here. I will argue that a general theory of licensing forces us to assume *right-adjunction* (rather than for example a right-branching analysis). Moreover we will see that Identification is the only licensing requirement that can be assumed for extraposed elements. Finally I will give a brief outline of the main assumptions of my analysis, the predictions it makes and the (traditional) problems it can solve.

Chapter II. to IV. are mainly intended to justify the proposal w.r.t. some apparent problems (i.e. to justify departures from standard assumptions). However in analyzing these apparent problems the analysis will derive some results that will turn out to be quite welcome (on basis of both empirical and theoretical considerations).

In Chapter II. I will start to show how the proposal works in detail. For doing so I need to make some additional assumptions, which are not quite standard. For justifying my claims I will on the one hand give some theoretical and empirical support that these assumptions are independently necessary. On the other hand I will also show that some similar assumptions are to be found in the literature, however in a slightly different way. I will argue that *domain D* has to be seen as an *independent level of representation*, thus my analysis will have some crucial implications for the assumption concerning the organization of Grammar. I will furthermore propose a particular analysis for the *representation of nominals*, which seems necessary for my analysis to work. It will be compared to other proposals found in the literature among

those found in a dynamic theory of meaning (Discours Representation Theory and file change semantics respectively). Throughout the following chapters this representation will be used and also shown to be independently motivated. Here I will also argue that it is necessary to assume sentential elements to be associated with a discourse referent and I will discuss the general issue of reference to abstract objects in discourse (in the sense of Asher (1993).

Chapter III. addresses one apparent problem of the proposal, i.e. I will deal with extraposed clauses that are correlated with a *pronominal* in the matrix clause (Notice that this is in general the main empirical domain I am looking at, i.e. extraposed elements which are construed with some (overt) antecedent). I will justify the necessary claim that pronominals can introduce a new discourse referent - although in a special way. However we will see that this assumption will derive some otherwise unexpected presuppositions. Some of the relevant data will also lead us to address the issue of optionality (which is crucially involved in extraposition). The assumptions made there will enable us to look at extraposition from the point of view of minimalistic assumptions (in the sense of Chomsky 1992, 1993, 1994). Domain D will be argued to be a crucial step towards analyzing some (apparent) optional phenomena. Moreover I will address the general issue of accessibility for anaphoric resolution concerning discourse as well as syntax. The phenomenon of *backward pronominalization* will be discussed. I will argue that the well known presuppositions arising in this array of empirical domain can be analyzed on par with extraposed clauses which are construed with a pronominal element. Moreover this analysis will also provide some indirect support for assuming domain D as an independent level of representation.

In Chapter IV. I will in particular analyze extraposed clauses which are construed with a *definite DP* in the matrix. Here I will argue that definite DPs can introduce discourse referents. Although this is not really a new assumption I will depart from traditional views. We will see that both assumptions to be found in the linguistic and philosophical literature of definite descriptions are somehow correct: I will show that the cases of Identification analysed here lead to the conclusion that a definite DP can introduce the *uniqueness interpretation* (in the sense of Russell 1905) and the identifier (for example a relative clause) provide *availability in discourse* (i.e. familiarity in the sense of Heim 1982). Thus both treatments of what the definite determiner contributes are argued to be right - in most environments they occur together. This will also lead us to a discussion of some environments where previous analysis suggested a specificity (or definiteness) effect to be at act. Here the analysis suggested here implies that at least some of those environments seem to induce a *familiarity effect*.

I will then discuss the difference between pronominals and nominals and show how different properties of the constructions under consideration follow from a crucial distinction between both of them (we will see however that there is not simple clearcut distinction rather I will argue that there is a relevant notion of *scalar* difference involved). These assumptions will then lead us to address the problem of split antecedents.

We will furthermore see (as expected) that similar presuppositions arise w.r.t. pronominal antecedents. Finally this will lead to a discussion of *(contrastive) focus*. Since we will see two different presuppositions arising depending on whether the noun or the determiner is (contrastively) stressed we will see indirect support for the representation of nominals on the one hand as well as for the relevance of Identification on the other hand.

This conclusion will then lead us to an analysis of *indefinite*, *predicative* and *quantified NPs* construed with extraposed clauses in chapter V. Here we will see that some previously unrecognized properties are predicted by the analysis.

In chapter VI. I will then turn to some syntactic predictions of the analysis proposed. This section will show how the present analysis can solve some of the well-known properties of extraposition without referring to construction specific statements. We will there see that making use of precedence rather than rightness is also empirically justified. Here the issue of the relevant adjunction site of extraposition will also be addressed and it will be shown to follow quite straightforward from the analysis suggested here. Finally we will once more address the issue of optionality of extraposition.

Then, in an excursion I will provide an application of the present analysis for a very restricted empirical domain, i.e. I will be concerned with DPs containing *einzig* ('only'). Here some problems will be addressed which have been only sketched in some of the previous chapters. *einzig* is in particular relevant for the present discussion since it focusses on the meaning we assume to be associated with (most of) the correlates, namely uniqueness.

In Chapter VII. I will address one apparent problem for the present analysis, namely the construction referred to as *(Contrastive) Left Dislocation*. I will provide an independently motivated analysis for this construction which will lead to the conclusion that the apparent problem is not really a problem but rather a prediction of the present analysis. This claim will also be justified by showing that Left Dislocation is really the

opposite side of the same coin of the licensing relation we assume for extraposed constituents. This will be seen especially on basis of the presuppositions which arise in case of Left Dislocation.

Then I will turn to some extensions of the analysis. First in Chapter VIII I will address the issue of *Result Clauses* as well as *Conditionals*, both w.r.t. extraposition as well as Left Dislocation. All the attested properties shown in the 'core cases' addressed in chapter III.-VII. will be shown to also hold for these instances as well as some other instances where a correlate of an extraposed clause is found in the matrix clause.

In the last chapter I will finally compare the presented analysis with some previous analyses of extraposition. It will be shown that the relevant empirical facts are treated differently within the different approaches to be found in the literature on extraposition. the differences mainly reflect the general assumptions about the organization of grammar made in the respective frameworks. Thus analyses of (or the possibilities for analyzing) extraposition phenomena have changed in ways which reflect the development of the Theory of Grammar within the Generative framework i.e. from Transformational Grammar, towards a Principles & Parameters framework up to some fairly recent proposals (the Minimalist Program (Chomsky 1992); the strictly right-branching-hypothesis (Kayne 1993); Haider 1993)). Especially the latter approach cannot be overlooked in a thesis which has extraposition as its main topic. Thus whereas in the first chapter we will see some theoretical force to assume right-adjunction in the last chapter I will also explicitly show some empirical arguments against the right-branching (and in favor of the right-adjunction analysis).

I. LICENSING AND EXTRAPOSITION

The core claim to be defended in this dissertation is first that Identification is a (primitive) licensing relation within syntactic theory and second that extraposed constituents are licensed as Identifiers. Syntactic relations can in general be viewed as licensing relations. I will start with a brief discussion on the theory of licensing in general. Here I will try to establish the main (common) properties of licensing relations. Concerning Identification I will assume that this relation has the same properties.

1. General considerations concerning Licensing

It has been a main goal of syntactic theory to find answers to the following questions:

1) How is an element licensed in the position it occurs in?
2) What licensing requirements does a particular element impose?
3) What locality constraints hold for a particular licensing relation?

One could ask why the questions in 1 and 2 are distinguished. The answer is as follows. Consider for example the following sentences:

(1) (a) *Peter saw [John]*
(b) *Peter saw [himself]*
(c) *Peter saw [Mary leaving]*

The bracketed constituents are all licensed as 'complements' of the verb *see*. This is how they are licensed within their position. However, the three constituents differ in some further licensing requirements they impose. The two NPs *[John]* and *[himself]* differ from the VP *[Mary leaving]* in that they need case, i.e. NPs are subject to the case filter (which is a licensing requirement for NPs). The two NPs however, differ again: the reflexive pronoun *himself* needs an antecedent, i.e. it needs to be bound. The R-expression *John* is not subject to such a requirement (to the contrary, it must not be bound). On basis of these examples it should be clear why the questions in 1 and 2 above are to be distinguished. Different constituents can be licensed in the same way w.r.t. the position they occur in however, they can differ w.r.t. some further licensing requirements. To distinguish the two licensing requirements I will use the terms *inherent* licensing as opposed to *relational* licensing.[1] Question 1 concerns inherent

licensing whereas question 2 is about relational licensing. I will furthermore assume that Chomsky's (1992) notion of *legitimate LF-object* corresponds (roughly) to inherently licensed elements. He distinguishes four legitimate LF-objects: heads, arguments, modifiers and operator-variable-chains.

Concerning relational licensing the following property is the one which will mainly concern us here. It has to do with the denoting properties of a given element. In (1b) we can assume that due to the lack of inherent referential or denoting properties anaphors need a licenser to determine their actual content. In some sense they need to be identified. The same condition is relevant for empty categories. Notice that it is assumed for example in Rizzi (1986) that there are two different licensing conditions for *pro*: 1) formal licensing and 2) identificational licensing. Similar assumptions have been made in various different versions of the *Empty Category Principle* (ECP). The distinction between *head government* and *antecedent government* (or binding in the sense of Cinque 1990) reflects the difference between formal (i.e. inherent) and identificational (i.e. relational) licensing. Thus Identification is not at all a new notion within syntactic theory.[2] However I will argue that even apparently fully denoting expressions can induce the need for some further licensing w.r.t. picking out the actual (discourse) referent.

The second relevant property which I assume for licensing relations in general is that all of them are *asymmetric* in nature. We only find elements which need to be licensed by an other element. In addition, the licensing relation has to obey some well defined locality constraints. It is never the case that an element HAS to license another element. (Let me mention just a few instances: NPs need to be assigned case under government or Spec-head agreement, polarity items need to be c-commanded by an appropriate licenser, anaphors need to be bound within their binding domain, etc.). I assume this asymmetry to be an inherent property of licensing relations. This amounts to saying that there are no mutual licensing relations. Whenever we find two elements X and Y which seem to be mutually dependent on each other, there must be two

[1] It should be noticed that inherent licensing can also be mediated by an other element (for example the constituents in (1a-c) are licensed by means of selection by V). The only element which does not need any further licensing by another element is presumably the root node of a matrix clause. In general we can think of X'-theory together with the projection principle as being the licensing relations that are relevant for inherent licensing.

[2] Within discourse similar licensing requirements are at play. Heim's (1982) Familiarity-Novelty Condition is intended to capture this insight: abstracting away from several problems, the difference between indefinite and definite NPs is argued to be as follows: definite NPs need some discourse antecedent whereas indefinites must not have one.

licensing relations involved: X is licensed by Y w.r.t. a relation R1 - and Y is licensed by X w.r.t. another relation R2.[3]

2. Extraposition and Licensing

Let me now turn to extraposition and how it relates to the general properties of licensing as discussed in the previous section. The main empirical domain I will look at is sentential constituents (which can occur in extraposed position) that are linked to a nominal element in the matrix clause. Examples are given below:

(2) Pron...CP

(a) *Peter hat [es]$_i$ geglaubt, ([daß Maria Bier trinkt]$_i$)*
Peter has it believed that Mary beer drinks

(b) *Peter hat [daran]$_i$ geglaubt, ([daß Maria Bier trinkt]$_i$)*
Peter has thereon believed that Mary beer drinks

(3) NP.....Relative clause

Peter hat [die Frau]$_i$ getroffen, ([die gerne Bier trinkt]$_i$)
Peter has the woman met, who likes beer drinking

(4) NP.....complement clause

Peter hat [das Argument]$_i$ vorgebracht, ([daß Maria gerne Bier trinkt]$_i$)
Peter has the argument presented that Mary likes beer drinking

Here we find either a pronominal element (*it* or a *ProPP*) construed with a sentential constituent (2), a relative clause (3) or a 'complement clause' of an NP (4). As illustrated by these sentences, all the relevant data involve a relation between two (overt) elements. However, as indicated by the brackets the occurrence of the extraposed constituent is (syntactically) optional.[4]

[3] An example for an apparent mutual relation would be the *subject-predicate* relation. But here we also find that there are two different licensing relations at play: on the one hand the subject depends on the predicate (or the V) in terms of theta assignment. On the other hand the predicate is dependent on the subject because it needs to be saturated. Similarly Operator-variable chains can be interpreted in this way. The variable needs to be bound, whereas according to Sportiche (1983) operators need to be mapped to an A-position in order to be inherently licensed.

[4] It has to be noticed that the occurrence of *it* is optional in case the sentential constituent is present. I will discuss this property in section III.3.

According to what we have said about licensing in general the relevant problems to be addressed are the following ones. On the one hand the nominal element in the matrix clause is inherently licensed (as a complement of the verb). Since the extraposed clause is apparently optional at first sight it is not evident that the nominal element is necessarily licensed by (or dependent on) the extraposed constituent. On the other hand consider the extraposed clause. For the moment I will just assume following Culicover & Rochemont (1990) (henceforth C&R) that these elements are *base-generated* in right-adjoined position.[5] Hence they do not straightforwardly fall within the class of elements that are inherently licensed.

The following properties of extraposition are general problems for any analysis of extraposition. As already mentioned in the introduction one of the main problems is that the obligatory rightness of extraposition does not follow from any independent principle. It has to be stipulated. The second (empirical) problem has to do with the fact that rightward relations (as extraposition) seem to be both more and less restricted than leftward relations. Extraposition is upward bounded (i.e. it cannot undergo successive cyclic movement) so that it is more restricted than leftward movement. However, an extraposed element can be linked to a position from which leftward movement is excluded. Thus, there is an empirical difference between left- and rightward relations. Apparently, all of the traditional analyses account for this by claiming that there is a genuine (primitive) difference between left- and rightward relations. Moreover in many of the analyses the adjunction site of extraposition appears to be quite stipulative (a topic I will discuss in more detail in Appendix I.). Finally the apparent optionality of extraposition on the one hand and the optionality of the occurrence of the extraposed element on the other hand is also a problem given the general attempt within syntactic theory to get rid of optionality.

[5] In Appendix I. and II. I will explicitly argue for such a view.

3. The Proposal

Concerning the empirical domain the main points of the analysis are the following ones.

I. Every nominal element is associated with two indices (relevant at domain D) as indicated below:

(5) $[_{DP}\ \mathrm{Det}_x\ [_{NP}\ \mathrm{N}_X]_X]_x$

N is associated with an index X corresponding to a set of possible discourse referents which is restricted by the content of N. D is associated with an index x corresponding to the actual discourse referent, which has to be a subset of X.[6]

If the nominal element under consideration does not have enough descriptive content to pick out the actual discourse referent (x) in the domain of discourse, it induces the need for further (relational) licensing - it needs to be identified. This is the second main claim of the analysis:

II. The relation between the nominal element and the extraposed clause is an instance of *Identification at domain D*.

For the cases under consideration I assume that Identification is a licensing relation at domain D. For syntactic wellformedness conditions as well as for the assignment of truth conditions the nominal elements need not be licensed any further. I interpret this fact as to indicate that S-structure as well as LF requirements are satisfied. However, it can be the case that the nominal element does not have enough descriptive content in order to pick out the relevant discourse referent. This means that the set of possible discourse referents needs further restriction. The sentential constituent provides enough descriptive content to actually pick x out of X. They act as *identifiers* for the nominal elements, which I will refer to as *identifyees*. Since this licensing requirement is related to discourse properties I assume that the relevant level for establishing this licensing requirement is domain D (as I will discuss in chapter II.). This notion of Identification implies the following assumption:

[6] This representation is reminiscent of Vergnaud & Zubizarreta (1992) on the one hand and Enç (1991) on the other hand. I will discuss these approaches in section II.2.1.

III. The identifyee introduces a discourse referent.

Therefore it follows from Heim's (1982) Novelty Condition that there is a precedence constraint as given below:

The Precedence Constraint on Identification (PCI)
The identifyee has to precede the identifier.
(where XP precedes YP iff it linearly precedes every member of YP).

Finally, as all licensing requirements Identification obeys a particular locality constraint. The relevant locality constraint is given below:

IV. Locality Constraint on Identification (LCI)
The identifyee has to be i(dentificational)-governed by the identifier.

i-government:
X i-governs Y if X c-commands Y (or X is a sister to Y) and there is no node *Z*, *Z* dominates Y and X c-commands Y.[7]

Assuming I.-IV. one immediately gets rid of the problems discussed in section 2. First and most crucially the obligatory rightness of the identifier is derived from the Novelty Condition (which implies the precedence constraint). Secondly the adjunction site of the identifier is derived by LCI. Thus, the identifyee determines the adjunction site. (This also captures upward boundedness as well as the Nesting Effect on multiple extraposed constituents). Other cases of 'extraposition' are neither subject to PCI nor to LCI. Thus there is no specified node for adjunction of extraposed constituents.[8] Thirdly the representation of nominals together with the properties of Identification at domain D will derive some pragmatic presuppositions, i.e. the sentences under consideration are not felicitous in every discourse context. Similar effects can be observed in case of backward pronominalization - and it will be shown that they follow from the same principle. These three predictions are crucially related in the following

[7] Notice that the exact definition of the LCI is not really crucial here as usual it depends on the definition of c-command. For the present formulation I assume the definition of c-command as given in Kayne (1993):
X c-commands Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y.
Descriptively LCI results in the following domain: the identifier can either be adjoined to the identifyee or to the first maximal projection dominating the identifyee.

[8] cf. Wiltschko (1994) for a more detailed discussion.

way. Rightness is replaced by precedence, a notion which is crucially related to linear ordering. We will moreover see that the S-structure position of the identifyee is crucial for determining the adjunction site. S-structure (rather than D-structure or LF) is input for the linear serialization of the relevant constituents. This means that LCI is related to linear ordering. Finally the presuppositions are crucially related to linear ordering as well, i.e. the S-structure position of constituents. Thus S-structure and its mapping to domain D is crucial for analyzing the different properties of extraposition (i.e. PCI, LCI and the presuppositions).

Finally, let me show how the analysis introduced above provides a solution for the licensing problem concerning extraposition. On the one hand the identifyee is relationally licensed by the identifier (the identifier provides descriptive content for the identifyee). On the other hand, for inherent licensing the identifier is dependent on the identifyee. The former appears in an A'-position and it is mapped to an element in an A-position (i.e. the identifyee) by the relation of Identification. We find a similar situation as in case of an Operator-variable chain where the Operator is mapped to the variable by means of binding. This means that the apparent symmetric (or mutual) relation of Identification is really an instance of two different licensing relations. The interaction of these two licensing relations necessarily results in obligatory right-adjunction: the identifyee has to precede the identifier but at the same time it must be governed by the identifier. Right-adjunction of the identifier is the only way to fulfill both constraints at the same time.

At this point, I will briefly address an important question within syntactic theory, namely the problem of how linear ordering and hierarchical structure interact. Within several earlier analyses (e.g. of binding facts) precedence has been argued to play an important role as a property of licensing requirements. However, it seems that one of the main goals of syntacticians was and is to get rid of the notion of precedence: all constraints, it is argued, should follow from hierarchical relations. Linear relations should just be a consequence. Although this seems to work for the notion of precedence in several cases, it didn't work for a related matter, namely directionality. Saying that a particular language is head-final or head-initial amounts to saying that within this language the head has to precede or follow its complement respectively. The same is true of the claim that a certain language assigns case to the right or to the left: this merely states that in a given language the verb assigns case either to the preceding or to the following element.

In his antisymmetry paper Kayne (1993) tries to get rid of all notions of linear ordering. He claims that linear ordering follows trivially from hierarchical structure.

In assuming this he solves the directionality problem because it follows from his analysis that possible phrase markers are universally head-initial. In this approach, c-command is not just a hierarchical relation which governs licensing requirements, but it is a (partial) ordering relation which maps into precedence.

A slightly different approach is found in Brody (1994). There it is assumed that dependency relations are mapped into precedence. In his approach hierarchical structure crucially involves dependency relations. All dependency relations are asymmetric and are from left to right (an element X which is dependent on an element Y is preceded by Y). This approach faces an important problem: in case of apparent mutual relations (where I have argued that we find two different relations) it has to be stipulated that only one of the two relations counts as a dependency relation. Otherwise no linear ordering between the two elements could be established. A comparable approach to Brody (1994), which does not face this problem, is adopted in Williams (1993). There it is assumed that 'Leftness' (i.e. precedence) governs dependency whereas coreference is governed by c-command. Notice that the present analysis is in the spirit of the latter assumption. The conclusion that the identifier has to be right-adjoined follows from two different relations: the identifier is dependent on the identifyee for inherent licensing (the former has to be preceded by the latter) whereas the identifyee is dependent on the identifier in terms of 'coreference': the denotational content is determined by the identifier. Therefore the identifyee has to be c-commanded/governed by the identifier. As opposed to Kayne (1993), I do not have to exclude right-adjunction as a possible option of UG. To the contrary: I am forced to assume right-adjunction. However, we can still stick to some notion of antisymmetry since the licensing relations involved are asymmetric.

II. IDENTIFICATION

Before discussing the predictions inherent to the analysis outlined in the previous chapter it needs to be clarified what I mean with *Identification at domain D*. Therefore I will discuss the notion of domain D as well as the representation of nominals I assume to be relevant for this level.

1. Domain D as an independent level of representation

Following Vergnaud & Zubizaretta (1992) (henceforth V&Z) as well as Prinzhorn (1992) I will assume that domain D is an independent level of representation. For justifying and explaining this proposal I will in this section briefly discuss some properties of levels of representations in general as well as some implications the assumption made here has for the theory of Grammar.

V&Z assume an additional level of sentence representation; *L-structure*, which they identify with domain D in the sense of Chomsky (1981: 324). This level of representation is crucially related to the interpretation of nominals. The way I am looking at domain D is slightly different. Domain D is taken to be relevant for capturing crucial properties of discourse. I assume that it is this level where well-formedness conditions are evaluated which are related to the *felicity* of a given utterance. The notion of domain D as used here will be shown to be related to Discourse Representation Theory (DRT).

Given the assumption of levels of representations, the question about the nature of this representation is a crucial one. The following possibilities arise: (a) it is a level of *logical form*, (b) it is a *partial model* of what is said in the discourse or (c) it is a *mental representation* of the content of a discourse formed by a recipient of it.

A representation which takes discourse into account (like domain D) is more related to the notion of *use* of language than a sole syntactic representation of a single sentence (since it is sensitive to what has been said before). The notion of "& use" however, was and is explicitly excluded within both *generative grammar* (or syntax in the narrower sense) and formal semantics. On the one hand the philosophical attempts to link *meaning* with *use* have been independent of formal semantics. On the other hand formal syntax explicitly excludes *language-use* from analyzing the form of

language. Syntax is assumed to be independent of pragmatic (the *autonomous hypothesis*). However there are two notions of *use* which need to be treated separately. On the one hand *use* can be interpreted in the sense of "real-embedding" in spoken language. Within generative theory this corresponds to the dichotomy *performance* vs. *competence*. The former being related to the 'use of language' (where several factors can intervene that are not to be taken into account in analyzing the language faculty). Competence on the other hand is assumed to be related to the human mind. Thus every human being (who does not suffer from any mental disability) develops a competence of her native language. Competence is the subject of linguistic research within the generative enterprise. The data to be analyzed are not taken from 'actual speech' or 'discourses'. What is important is wellformedness judgements of native speakers.

The other side of the coin is to interpret "use" of language within the range of (pragmatic) *competence*.[1] However, I do not think that domain D is only related to *pragmatic* competence. I think that this level of representation is within the scope of grammar in the narrow sense. The difference to a "strict" autonomous hypothesis of syntax is that it cannot be assumed that the only item of interest is a single constituent, i.e. a sentence. I assume Domain D to be a level of representation that is associated with a given sentence, but its well-formedness cannot be judged without taking other (previous) representations into account. To exemplify this point let us look at the *Familiarity-Novelty Condition* (in the sense of Heim 1982), which I take to be a well-formedness condition. On the basis of a single sentence one might be able to evaluate it (for example when a given NP occurs twice within a single sentence) but not necessarily so. If a given NP occurs just once in a given sentence then the discourse referent denoted by this NP could have been introduced in a previous sentence. Judging the well-formedness of single sentences w.r.t. well-formedness conditions that hold for bigger units is just a subset of the domain of evaluation (i.e. discourse). Since a sentence is a subset of the relevant domain of evaluation there are various interactions between the levels of representation relevant for pure syntactic phenomena and the level of representation relevant for discourse phenomena.

The same point can be made for certain *presuppositions* that I take to be a consequence of a particular representation at domain D. Here it has to be noticed that throughout this thesis I will use the term presupposition for capturing *pragmatic presuppositions* in the following sense: certain properties of a situation (or discourse) can impose conditions on the appropriateness of the use of a sentence S rather than a

[1] Notice that there are several attempts to analyze the nature of *pragmatic competence* (cf. Sperber & Wilson 1986).

necessary condition for the proposition expressed by S to have a truth value. The latter condition would correspond to *semantic presupposition* (presumably to be captured at LF). Pragmatic presuppositions are to be seen as relations between an utterance and a proposition whereas semantic presuppositions are relations between two propositions.[2] The notion of pragmatic presupposition has also received an interpretation which is related to the above mentioned Familiarity-Novelty Condition. Karttunen (1974) assumes that an utterance U presupposes a proposition P if that utterance is acceptable only at a point in a discourse where that proposition is in the set of propositions that the parties to the discourse take as established. We will see that both notions of pragmatic presupposition coincide in the cases we are interested here. The previous remarks should however support the claim that domain D is an independent level of representation.

Furthermore I take domain D as a level or representation which corresponds to a mental representation. Mental representations serve as objects for the human mind to operate with. Although they are created in an autonomous manner they have to be within an input- output-relation to the central system. The generation of such a representation is however assumed to be independent of the central system. It is autonomous.[3]

Restrictions applying at the level of domain D are neither clause-bound nor sentence-bound. They are relevant for the domain of discourse. Therefore I take domain D to be a different representation to the level of logical form (LF). The latter is more related to formal semantics. It should provide a representation which suffices to determine the interpretation and the truth conditions of a single sentence (i.e. scope properties, binding, etc.). Domain D is related to (formal) pragmatic. I do not believe that the two levels can be reduced to each other[4], nor do I believe that there is a

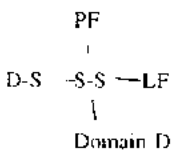
[2] Cf. MacCawley (1991) for a discussion on the distinction between the two notions.

[3] Notice that Chater, Pickering & Milward (1994) draw a similar conclusion on basis of psycholinguistic evidence:
"We suggest that a more attractive approach is to separate the representation used for compositional semantics from that used in inference" (Chater, Pickering & Milward 1994: 11)
This suggests that we are really dealing with a separated level of representation corresponding to a mental representation.

[4] Notice that such a view is taken in Prinzhorn (1992). He assumes the following 'sandwich model': D-S - S-S - Domain D.

'derivational' input-output-relation between the two levels.[5] I assume that these levels of representation are independent of each other. This means that we have to reinterpret the standard T-model as follows:

(1)



Here it is crucial that S-structure branches into four independent levels (D-structure PF, LF and domain D) all three being independent from each other.[6]

1.1. Domain D and Minimalism

The above mentioned T-model has recently been suspended with in the minimalist program as introduced in Chomsky (1992). Let me briefly mention how the argumentation there is relevant within the present analysis[7]. Chomsky argues that Grammar is to be taken as a system consisting of the lexicon and a computational system. He dispenses with D-structure as an independent level of representation. In order to doing so he reintroduces *generalized transformations*, claiming that insertion need not take place in an "all-at-once" manner (as implied by the assumption of D-structure, relabelled as SATISFY). Items are taken out of the lexicon, structure is created and at one point (i.e. SPELL OUT) the computation branches. After SPELL OUT access to the lexicon is blocked. Then computation upon a single phrasemarker

[5] Under this view domain D would be similar to LF' in the sense of Safir (1986). Cf. Vallduví (1990) for a discussion of *Informational Structure* as neither being derived by LF nor to be equated with LF.

[6] cf. Vallduví (1990) for a similar conclusion concerning the representation of 'informational structure':
"Chomsky (1988) views S-structure as a 'contact' level between fundamental levels of representation: 'From this standpoint S-structure is a derived concept. For a specific language L, its properties are determined by those of the fundamental levels, and the condition that it be related to them by the appropriate principles' (1988:3). Each of the 'fundamental levels' is the structural interface with an autonomous linguistic component, and each sentence is a bundle of 'fundamental' abstract representations that must all be reflected somehow at S-structure to allow for the interface with the physical reality of an utterance" (Vallduví, 1990: 165)

[7] Notice that throughout this thesis I will adopt my the Principles and Parameters framework.

can continue. The picture there is as follows[8]:

```
(2)    Lexicon — SPELL OUT ——> LF
                      \
                       PF
```

S-structure as an independent level of representation is also suspended with. In a broad way corresponds to the point in the derivation where SPELL OUT takes place. The difference is that this point is not taken to be an independent level of representation where syntactic restrictions or well-formedness conditions apply.

Chomsky's line of argument is that the level of PF and the level of LF are independently justified since grammar is embedded in two *performance systems*: one corresponds to *articulation* and *perception* and the other corresponds to a *conceptual-interpretational* component. These two components have access to the phrase marker created by the computational system. It is argued that they are necessary on "virtual conceptual necessity". The two (remaining) levels of representation are "interface - levels". The minimal assumption then is that besides these independently necessary levels of representation no others are necessary and therefore they do not exist.

If it can be shown that certain restrictions apply at a level of representation, then one has good reason for assuming this level as an independent level of representation. The strongest argument for the existence of such a level is of course to show that this restrictions can only hold at this level and not at any of the other (independently motivated) levels. Consequently Chomsky has to show how the various restrictions previously assumed to hold at D-structure or S-structure respectively can be suspended with. One has to assume that they hold at a different level of representation or they

---

[8] Notice that in Chomsky (1994) the question as to where representations having to do with discourse are to be found is briefly sketched. He assumes a line of reasoning which is in the spirit of the earlier EST framework, where so called (optional) 'stylistic' rules were taken to be a pure PF-phenomenon.
"Notice that we are sweeping under the rug questions of considerable significance, notably, questions about what in the earlier EST framework were called "surface effects" on interpretation. These are manifold, involving topic-focus and theme-rheme structures, figure-ground properties, effects of adjacency and linearity, and many others. Prima facie, they seem to involve some additional level or levels internal to the phonological component, post-morphology but pre-phonetic, accessed at the interface along with PF and LF." (Chomsky 1994: 1f.)
Some of these phenomena are relevant for domain D. I will briefly discuss the issue of 'extraposition as a PF-phenomenon' in Chapter VI. There we will see some crucial arguments against such a view. Moreover I refer the reader to Guéron (1980) where it is convincingly argued that extraposition has to be seen as a 'rule of core grammar'.

have to be defined as restrictions on the computational system itself.

I think that both lines of reasoning to be found in Chomsky (1992) are crucial for the assumption of domain D as an independent level of representation. Once you can show that a certain restriction can only apply at domain D then you can conclude that this level exists. Consider under this light the Familiarty-Novelty Condition. It is a restriction that applies to bigger units as the mere sentence-level. This means that it is a restriction which does not only hold for the output of the computational system as defined above. This can be taken as an argument that the Familiarity-Novelty Condition cannot be assumed to hold at one of the levels that are considered as the output of the computational system (generating a single sentence). Still this representation has to be associated with a given sentence.

I think that domain D is also motivated by the second line of reasoning Chomsky takes. One can assume that domain D is the interface level to 'pragmatic competence'. This level of representation is assumed to stand in an input-output relation with a component which operates neither on truth-value assignment nor on articulatory or perceptual properties but on discourse properties. This is to say that the felicity or appropriateness of a given sentence in a certain (discourse) environment is to be captured at this independent level. We will see how certain presuppositions can be captured at this level. If a sentence is associated with a given presupposition then it is just felicitous in a certain context but ill-formed in a context which does not allow for such a presupposition.

Assuming the possibility to talk about domain D within the minimalist framework I would like to speculate a bit more on this topic. Consider again Chomsky's (1992) assumptions. His claim is that all languages have the same LF-representations. (One of the reasons for this assumption is that parametric differences should be detectable on basis of overt (morphological) properties in order to explain language acquisition). Languages can however differ w.r.t. the "position" of SPELL OUT in the course of the computation. The claim is that if a given movement is detected within one language then every language has this movement - if not overtly then covertly (i.e. after SPELL OUT). Now the question as to why movement has to take place gets an additional dimension: what triggers overt movement in a given language A whereas it can take place covertly in another language B. As a solution Chomsky assumes that overt movement is triggered by the presence of a "strong feature" which has to be deleted (by means of checking this feature) before the derivation branches to SPELL OUT (strong features are argued to be illegitimate objects at PF). If another language B has the corresponding "weak feature" then

movement can and therefore (by means of economy) has to be delayed until "LF" (covert movement being less costly then overt movement).

To exemplify this claim consider the well known difference of the position of the verb in French and English (originally discussed in Emonds 1970):

(3) (a) *Peter often kisses Mary*
(b) *Peter aimer souvent Marie.*

According to the minimalist program this distinction reduces to a distinction of verbal features: in English the verb has weak features and therefore it can stay in situ until SPELL OUT has taken place. In French on the other hand the verbal features are assumed to be strong triggering overt movement.

Consider now a discourse related phenomenon like focus. What one could say within the minimalist framework is that a feature [focus] or [topic] is assigned to a given constituent. Languages might differ as to whether this feature is "strong" or "weak" respectively. If it is strong the constituent associated with it has to move before SPELL OUT. If it corresponds to a weak feature it can survive at PF and movement would take place covertly. This could for example explain the difference between Hungarian and English respectively:

(4) (a) *Mari [$_F$ JANOST] látta a kertben*
Mary John-acc 3s-pst-see the garden-in
(b) *Mary saw JOHN in the garden*

In Hungarian there is obligatory movement of a focussed constituent to a specified position whereas in English focus is indicated by means of stress on the relevant constituent which is still in its base-position. Notice moreover that it has been assumed (for example in Chomsky (1980) that stressed (i.e. focussed) constituents are also moved covertly in English, to account for some apparent weak cross over violations as in the following example:[9]

(5) **You deal with me and his$_i$ parents will deal with JOHN$_i$*

[9] cf. section III.4. for a discussion of similar data. There I will assume (following Williams 1994a) that the ungrammaticality can be reduced to some principles governing the occurrence of backward pronominalization.

However, I think that such an approach faces some serious problems. The distinction between strong and weak verbal features should be reducible to morphological properties in order to avoid circularity. But a feature like topic or focus is not a matter of pure morphological properties (although it is true that some languages mark this feature morphologically, as for example Japanese or Navajo). A feature like this would more or less correspond to the wh-feature. For the wh-feature Chomsky (1992) assumes that it is invariantly "strong". It is therefore assumed that all languages have overt movement of the wh-operator. Some languages make use of a silent wh-operator others pied-pipe the lexical wh-element with the empty operator. However, if this is assumed then the original motivation for postulating strong and weak features (on basis of the difference of overt word order) is undermined, since a feature can be strong but nevertheless movement is not visible (because what moves overtly is a silent category). I think that the same problem arises if one postulates a topic of focus feature.

Moreover it has to be assumed that features like 'topic' and 'focus' are primitives within the theory of grammar. As opposed to morphological features it cannot be assumed that they come along with an item from the lexicon. They have to be assigned in the course of the computation. If so then it is questionable whether language variation can still be reduced to lexical properties. I believe that it is more in the spirit of a Principles and Parameters framework to assume that the interpretation as topic or focus is a "result" of a certain configuration or process rather than a primitive feature.

The second problem I see with the minimalist assumption 'that all languages have the same LF-representation' is the following one. If a language has some sort of overt movement process then so it is argued all languages have this movement (at least covertly). I think the danger of such an assumption is to fall back into construction specific statements. Consider for example German relative clauses. The relative pronoun moves to SpecCP overtly. If movement is reduced to the need of checking some given features then the pronoun would have to carry a feature to be checked in SpecCP. This feature should moreover be a morphological feature. Thus we would expect these features to be present no matter where the pronoun occurs. Thus it is not at all obvious why pronouns of the very same shape as relative pronouns (i.e. demonstrative pronouns) do not always move to SpecCP.[10] The question arises as to why it is possible to have both word orders below in one and the same language:

[10] Under such an assumption the fact that in many languages the relative pronoun corresponds to the demonstrative pronoun would be a mere accident.

(6) (a) *Ich habe den gesehen*
I have pron seen
(b) *den ich gesehen habe.*
whom I seen have

It seems necessary to assign features to a given element in a way which is sensitive to the "construction" it occurs in. This amounts to going back to construction specific rules. Thus it is just in comparing different sorts of constructions that we can compare different languages and the differences w.r.t. overt or covert movement. But comparing constructions is not at all in the spirit of an approach which takes a construction to be the output of the interaction of various interacting principles and parameters.

2. The implementation and its possible translations

For representing a certain phenomena one is in the need to formally implement the relevant properties and objects which are assumed to be primitives, i.e. not reducible. In general the decision for a certain implementation is in many cases not just an empirical matter, but sometimes there are also theory dependent (or framework-dependent) guidelines. The choice of how to choose the adequate implementation is also a question of how fine-grained the analysis should be. Sometimes it suffices to abstract away from certain facts (which are by no means irrelevant for some other phenomena). In choosing an appropriate implementation for an analysis one needs to find a way between being fine-grained enough to capture all the relevant facts under consideration but also general enough to abstract away from other facts that do not seem to be relevant. I will first introduce the representation of nominals used throughout this thesis. For completeness I will also show how this implementation translates into other proposals found in the literature (V&Z, Stowell 1987, Enç 1991 and DRT).

2.1. The representation of nominals

I will assume that every nominal element is associated with two different indices as indicated below:

(7) $[_{DP}\ [_{D'}\text{the}]_x\ [_{NP}\ [_{N'}\text{man}]_X]_X]_x$
NP: X = set of individuals satisfying the property denoted in N
DP: x = the referent (in domain D)
$x \subseteq X$[11]

X is introduced by N (and thus it is associated with NP). It denotes the set of individuals satisfying the property denoted by N. It has to be noted that the denotation of X has to be relativized to the *universe of discourse*. x is introduced by D (and is therefore the index associated with DP). It denotes the actual discourse referent, which also satisfies the property denoted in N. Since I assume the set denoted by X is defined extensionally (although it really denotes a property) we always find a subset relation: $x \subseteq X$. This part of the analysis will turn out to be crucial. By means of this representation we can derive some crucial presuppositions associated with the interpretation of the relevant nominals. I will assume that this representation (relevant for domain D) is associated with all sorts of different nominal elements (including Names and pronouns). The differences found among the different nominals should thus follow from various properties of the elements under consideration, some of which we will discuss more carefully once we proceed. Moreover I will also assume (differing from other proposals) that similar facts hold for non-nominal elements. Thus I will differ from V&Z and DRT respectively in assuming that sentential elements are also associated with a discourse referent.

2.1.1. Vergnaud & Zubizaretta (1992)

The assumption introduced in the last sections is much in the spirit of V&Z. As I have already mentioned it is assumed there that domain D is an independent level of representation (referred to as L-structure). L-structure is defined as follows:

(8) *L-Structure is a set of elements D = {i,j,k,...} such that each element of D is associated with some nominal phrase in the sentence.* (V&Z: 610)

[11] I assume $x \subseteq X$ rather than $x \in X$ in order to capture plural (definite) DPs. In most of the cases x will only denote a singleton set. Thus the present representation involves an abbreviation. We should put it as $\{x\} \subseteq \{X\}$

The elements of L-structure are referred to as *denotata*. L-structure functions as an indexing device of nominal categories. It is crucial that the notion of denotation is independent of the notion of *reference*. Referential elements are included in the notion of denoting elements. V&Z assume that:

> *"Grammar does not include any autonomous level of representation that would be the set of entities referred to by the nominal expressions in the sentence."*(V&Z: Fn. 23)

Notice that this assumption is much in the spirit of Karttunen (1971) where the concept of *discourse referent* was introduced as a mental construct. It is assumed to be a mediation between referring expressions and real world referents. V&Z assume that an L-structure (i.e. domain D) representation that is associated with a given sentence is a set of indices (10a). Furthermore there is an additional mapping between the nominal expressions and the corresponding denotata (given in (10b):

(9) *The children like Disneyland*

(10) (a) {i,j}

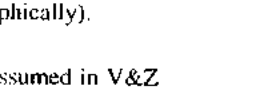
(b) {([$_{NP}$ the children],i), ([$_{NP}$ Disneyland],j)}. (V&Z: 610)

According to this representation *the children* in (9a) denotes i and *Disneyland* denotes j. Therefore:

> *"given some nominal expression XP in a sentence s such that XP is associated with the denotatum n in the L-structure representation of s we say that "XP denotes n".*
> (V&Z: 610)

One of the empirical facts V&Z capture with their analysis is the difference between a *type* and a *token* interpretation. The distinction between types and tokens is assumed to be a primitive distinction within grammatical theory that is represented at domain D (and it corresponds to a mental distinction). The idea is implemented as follows. They assume that there are two different indices: numbers (1,2,3,...) to designate tokens (and associated with DPs) and lower-case letters (i,j,k...) to designate types. Therefore the following representation for the definite DP *the cat* is assumed:

(11)



(V&Z: 613)

The (complement) relation between D and NP is assumed to be the projection of the instantiation relation that holds at domain D between the token 1 denoted by DP and

the type j denoted by NP. Moreover they assume that a so called *Correspondence Law* governs the relation between syntactic categories and semantic types in the case of non-pronominal definite nominal expressions:

(12) Correspondence Law
When a DP or an NP denote the DP denotes a token and the NP denotes a type. (V&Z: 612)

Any principle that is supposed to regulate the relation between semantic and syntactic categories faces various difficulties. There are two competing approaches concerning this question. They can be summarized as follows. On the one hand Russell's theory of definite descriptions claims that there is no correspondence between syntactic and logical constituency. On the other hand Montague (1970) argues that there is such a correspondence:

> *"There is a uniform correspondence between the categories of English and the types of IL (intensional Logic): that is, if an English expression of category x must translate into an IL expression of type a, then all English expressions of category x must translate into expressions of type a."* (Montague 1970)

It is assumed that a distinction between two logical categories must correspond to a syntactic distinction but the reverse is not true i.e. it is not assumed that there is a one to one mapping between syntactic and semantic categories. For example both common nouns and intransitive verbs translate into one-place predicates. However Montague still assumes the possibility of direct interpretation, (i.e. rules from semantic interpretation must exactly follow the syntactic rules of phrasal construction).

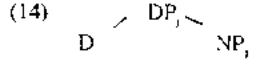
Williams (1983) takes some 'in-between-view' claiming that there is a strong correspondence between surface constituency and logical constituency, which does however not identify the two. There it is convincingly argued that one cannot assume a simple one to one mapping between syntactic categories and the interpretation they receive. One instance would be the distinction between referential and predicative NPs: under standard assumptions they are of the same syntactic category but correspond to a different logical type. This claim does not deny the relevance of a semantic type, it simply states that it can be instantiated in different ways (i.e. isomorphically).

The problem a strict correspondence law for DPs and NPs as assumed in V&Z has to face is circumvented there by introducing the notion of *expletiveness* at domain D. This means that syntactically there can be a constituent of category DP but it is not

represented at domain D i.e., it is expletive w.r.t. domain D. The following example is analyzed as involving such an expletive determiner. This has the result that syntactically a DP is present, but it is not interpreted as a token:

(13) *On a donne le meme ordinateur a Sophie, a Justine, et a Clea.* (V&Z: 614)

Sentence (13) is ambiguous between a token and a type interpretation. Under the former there is *one particular computer Sophie, Justine and Clea have got*. Under the type reading each of the individuals mentioned has got its own computer. The second reading apparently contradicts the correspondence law: there is a DP which does not denote a token but a type. The definite determiner does not seem to play a role for the interpretation. V&Z analyze this phenomenon by claiming that the definite determiner does not bear any denotational index. The determiner functions as an expletive from the point of view of denotation. It receives the following representation:

(14)


In this case they assume that the complement relation between D and NP is licensed morphologically (rather than as a projection of the instantiation relation). The presence of D is licensed by means of the overt person, number and gender agreement that holds between the determiner and ist complement NP.

The implementation I have chosen for the present analysis is in the spirit of V&Z. that corresponds to their type interpretation is the denotation of NPs: X (i.e. properties defined extensionally) . Their token interpretation corresponds to x. The relation of instantiation is somehow reminiscent to the subset relation I argued to hold between x and X. We will furthermore see that the notion of expletiveness for domain D will become important for some instances of apparent optionality (cf. section III.3). In addition I will argue that besides the type-token distinction (redefined in the present analysis) certain presuppositions are also to be captured at domain D.

2.1.2. Stowell (1987)

Stowell (1987) assumes that an NP is a predicate expressing a property thus providing a variable which needs to be bound. A determiner has the ability to bind this variable

thus saturating the predicate.[12] An NP is assumed to be an (unsaturated) non-referential predicate and a DP is a assumed to be a (saturated) referential argument. It is assumed that the distinction between predicates and arguments cuts across syntactic categories. Therefore Stowell (1987) faces the same problems as V&Z. It is predicted that every DP is a referential argument. As is well known there are nominals (containing a determiner) that are still predicates. Again there are two ways out. One can stick to the claim that there is a one to one mapping between syntactic and semantic categories. This is the line of reasoning Stowell (1987) advocates. According to this assumption every predicative nominal would be an NP containing a determiner that occupies SpecNP. The second possibility would be to say that determiners just have the ability to bind the variable provided by NP but (some of them) do not necessarily do so.[13] This would be similar to V&Z's claim that a determiner can act as an expletive (in their account from the point of view of denotation, in a binding account w.r.t. binding the variable). Both approaches face the problem as to how the determiner itself is licensed. (Remember that V&Z solved this problem by claiming that the an expletive determiner is morphologically licensed). The relevant question here has to do with predicational contexts. It is assumed that the variable provided by the NP is bound by the subject of the predicate. Although it is possible for some Ns to occur (optionally) without a determiner it is not true for all Ns and also not in all environments:

(15) (a) *Anne's death made the elector of Hannover (the) king of England*
(b) *Mary considers Bill (the) guardian of her children*
(c) *The students elected John president of the class*

(16) (a) *I found Jones *(the) only reliable hatmaker in town*
(b) *Mary considers Bill *(a) fool*
(c) *This book will make John *(the) most famous person I know*

(Stowell 1987)

(17) (a) *Peter is *(a) teacher*
(b) *Peter is *(an) idiot*

Moreover there are crosslinguistic differences concerning the obligatory vs. optional presence of the determiner respectively. German does not allow a bare NP in the same environment where English does. However German (as opposed to English) allows bare NPs in copula constructions with some (profession-denoting) NPs:

[12] This is also in accordance with Higginbotham (1985).

[13] Notice that this line of reasoning would be in the spirit of the assumption concerning licensing relations as discussed in section I.1.. There are no licensing relations where an element X HAS to license another element Y, i.e. it cannot be the case that the determiner has to bind the variable, but it is just the case that the variable has to be bound.

correspondence problem' is to assume that DPs are always referential. However the presence of a determiner is not a reliable test for referentiality. The determiner can also be adjoined to N' and therefore does not project a further DP. V&Z on the other hand claim that in the presence of a determiner a DP is always projected. However, the determiner can be expletive from the point of view of denotation. Here I will adopt an analysis in the spirit of V&Z's. I will return to this issue in section III.3.. There we will see that some of the phenomena under consideration as well as the problem of optionality receives an interesting interpretation under such a view.

2.1.3. Enç (1991)

Another way of implementing the representation of nominals is adopted in Enç (1991), for analyzing the semantics of specificity. Since there are some similarities to the analysis presented here I will briefly discuss her approach.

According to Enç the difference in specificity of a given NP correlates with a difference in the domain of discourse in which a given sentence is wellformed. The difference in meaning does not lie in the assignment of truth conditions but in properties of NPs that structure the domain of discourse. Enç also assumes the relevance of discourse as opposed to simply analyzing single sentences. Therefore her approach is rather similar to what I am assuming here. Moreover she assumes that the notion of specificity is strongly related to the notion of definiteness on the one hand and Pesetsky's (1987) notion of D-linking on the other hand. Pesetsky notes that wh-phrases of the form '*which N*' are just wellformed if the individuals found in the answer stem from a selected set of the domain of discourse. D-linking is the very notion Enç assumes to be relevant for specificity.

Enç assumes that specific NPs are semantically interpreted as partitives. Notice that this is reminiscent of the subset relation assumed to hold for the relation between x and X). Therefore she assumes that specificity imposes a constraint on the structure of domain D in addition to the constraint which is imposed by definiteness. This is implemented by extending Heim's (1982) theory of definiteness to specificity along the following lines: Every NP bears a pair of indices. Both indices carry a feature for definiteness; the feature on the first index determines the definiteness of the NP. The feature on the second index determines the specificity of the NP by means of constraining the relation between the referent of the NP to other discourse referents in the following way:

(18) (a) *Annas Tod machte den Fürst von Hannover zum[14]/*zu König von England*
Ann's death made the elector of Hannover to-the/to king of England
(b) *Maria hält Hans für *(denleinen) Beschützer ihrer Kinder*
Mary considers Hans for (the/a) guardian her$_{GEN}$ children
(c) *Die Studenten wählten Hans zum/*zu Vorsitzenden*
The students elected Hans to-the/to president

(19) (a) *Peter ist (ein/der) Lehrer*
Peter is (a/the) teacher
(b) *Peter ist *(ein/der) Idiot*
Peter is (a/the) idiot

Stowell (1987) addresses this question and resolves it by claiming that determiners can have two distinct functions: they are either generated in D^0 binding the variable provided by N. In doing so they turn the predicate into a (referential) argument. However, determiners can act as operators. In this case Stowell assumes that they are adjoined to N'. If so they are analyzed as converting a predicate of *kind instantiation* into a predicate of *kind membership* (defining a function from instantiation to membership). He claims that N's denote *kind sets*. (Notice that my analysis claiming that N denotes a set of individuals satisfying the property denoted by N is somehow reminiscent of such an analysis). If the set is a singleton set (like in case of *president* or *king*) a determiner is not necessary. It is assumed to be reminiscent of adjectives, which can also denote a property directly. However if N denotes a non-singleton set a determiner is necessary in order to denote a single member of this kind-set. However, as the examples above show this cannot be true for German since e.g. *teacher* does not denote a singleton kind-set. Thus Stowell's (1987) way of overcoming the 'strict

[14] '*zum*' is a preposition with an incorporated determiner:
(i) *Hans geht zu dem/zum Mann*
John goes to the/to-the man
It has to be noticed that the preposition with an incorporated determiner cannot occur in the same environment as the non-incorporated counterpart. In particular, a predicative context does not allow for the full lexical determiner:
(ii) **Ihr Tod macht ihn zu dem König.*
Her death made her to the king
This restriction might have to do with the fact that a non-incorporated determiner is necessarily (interpreted as) stressed, and stressed determiners need a restrictive relative. Thus compare:
(iii) *Ich habe DEN Mann gesehen, *(der Maria geküßt hat)*
I have THE man seen *(who Mary kissed has)
(iv) *Ihr Tod macht ihn zu dem König, der er jetzt ist*
Her death made him to the king, who he now is
(cf. chapter III. for a discussion of the issue of optionality and chapter IV for a discussion of stressed determiners.)

(20) Every $[_{NP}\ \alpha]_{<e,j>}$ is interpreted as $\alpha(x_j)$ and
$x_i \subseteq x_j$ if $NP_{<e,j>}$ is plural.
$\{x_i\} \subseteq x_j$ if $NP_{<e,j>}$ is singular. (Enç 1991: 7)

According to (20) an NP is specific iff its second index is definite. If the index j is definite the familiarity condition requires that the discourse referent x_j has already been established in domain D. Since (20) requires that the referent of NP is a subset of x_j the NP receives a specific interpretation. If the index j is indefinite it has to obey the Novelty Condition and the previous discourse cannot contain a referent x_j. A new referent is introduced within the domain. This referent has x_i (the referent of NP) as its subset. Enç notes that the properties of the novel superset are just those recovered from the head noun of the NP. Notice that this is again similar to what I assume here. The difference however is that I am assuming that NP and DP are associated with a single index each and by this they denote different entities. NP denotes X, which corresponds to a property (defined extensionally as a set) and DP denotes an individual. In assuming this we also get the result that N restricts the set out of which the actual discourse referent is taken. This results in the subset relation $x \subseteq X$ which obtains within a single DP and not only between two nominals (as it is the case in Enç's approach).

2.1.4. Nominals within Discourse representation theory (DRT)

Finally, I will now show that the implementation used here is translatable into a DR-theoretic framework. However, it seems to me that there is one crucial property which is missing in a straightforward and direct translation to a DR-theoretic account. This has to do with the subset-relation (i.e. $x \subseteq X$). We will see that this subset relation derives some crucial presuppositions associated with the examples under consideration. This part of the analysis is not easily available within DRT. Notice that DRT analyzes the meaning of a sentence in terms of a function from contexts to contexts. I think that such an approach should be able to derive these pragmatic presuppositions (i.e. felicity conditions imposed by certain contexts). In the course of the subsequent discussion I will compare the present analysis to DRT in case my account contradicts standard assumptions found in DRT.

Given the assumptions concerning the interpretation of DPs let us see how this works within DRT. In general DRT creates discourse representation structures (DRS) which are then compositionally construed to a full-representation that counts as the input of interpretation. Thus sentence (21a) is associated with the DRS in (21b):

(21) (a) *A boy kicked Fred.*

(b)
x,y
boy (x)
Fred (y)
kick (x,y)

Such a DRS consists of two parts: the first part is called the *universe* of the DRS the other is the *condition set*. This representation can also be represented as an ordered pair: $<U_K, Con_K>$. The universe contains discourse individuals (x and y). The conditions are formed from n-place DRS *predicates*, which take the discourse referents as their arguments. This has to be seen as a collection of property ascriptions to x and y. A DRS is compositionally construed out of some smaller units of DRS. DRS predicates can be represented as DRS as well (i.e. *predicative DRSs*).[15] A predicative DRS corresponding to *man* translates as:

(22) λx
man (x)

λ indicates an operation of discourse referent abstraction in which a discourse referent is removed from a DRS. Thus predicative DRS can combine with one or more discourse referents to yield a DRS. What about the determiner? A determiner according to DRT introduces a partial DRS, this is what remains from a DRS when a predicative DRS is abstracted away from it - thus partial DRS introduce discourse referents. A determiner is a function form predicative DRS to DRS.

In order to implement some version of the Familiarity-Novelty Condition the translation of a determiner must be relative to a DRS context. For example for an indefinite DP to be felicitous one has to make sure that the discourse referent it induces is not already present in U_K. Thus the translation of a determiner like English *a* is as follows (where u corresponds to a new discourse referent):

(23) λPλQ
u
P (u)
Q (u)

[15] For completeness let me give the definition of predicative DRS (taken from Asher 1993):
Definition: For a sequence of dr variables $x_1,...x_n$ and DRS K such that $x_1,...x_n$ element of U_K, $\lambda x_1....\lambda x_n<(U_K-\{x_1,....x_n\}), Con_K(x_1/x_1,....x_n/x_n)>$ is a *predicative DRS*. (Asher 1993: 70)

In this notation P and Q combine with predicative DRSs while u is a declared discourse referent. Now we can combine the partial DRS introduced by the determiner and the predicative DRS introduced by NP to get the following partial DRS (which can furthermore be combined with another predicative DRS for example the one associated with a VP):

(24) λQ [u | man (u), Q(u)]

Now let us compare this approach to my analysis. The DRT - version of claiming that an NP induces an index (X) which denotes the set of individuals satisfying the property denoted by N is to say that it yields a predicative DRS, which is an ascription of a property to a certain (given or introduced) object. Within DRT a property (P) is assigned to an individual. This then yields the discourse referent. In our implementation this translates as follows: out of the set of individuals satisfying the property P one (particular) individual is picked out. Since the determiner introduces another discourse referent (x) which is a subset of X the determiner can be assumed to be responsible for picking out the actual discourse referent. It is not obvious how DRT can implement the subset relation I am proposing. It is just the case that the determiner closes an open position within a predicative NP (by means of introducing a discourse referent which the predicate can take as its argument).

Notice that the DRT approach is reminiscent of Stowell's (1987) and Higginbotham's (1985) binding approach. It might well be that the binding- (saturation) approach is the relevant representation of nominals at LF. It is necessary for assigning truth-values. On the other hand the approach I am assuming can be taken to be only relevant for domain D, where discourse properties are relevant. If true this might give us an interesting solution concerning the licensing problem of determiners. In addition to being morphologically licensed it could be assumed that the determiner must either be represented at LF (i.e. binding the variable of N) or it must be represented at domain D (resulting in a token interpretation). For the type (distributive inalienable) interpretation discussed in V&Z this would mean that the determiner still binds the variable provided by N. There is a way to test this assumption. Suppose we find an instance where the definite determiner cannot be expletive w.r.t. domain D but has to be expletive w.r.t. LF. The relevant data to look at are predicative constructions (where the determiner is assumed to not bind the variable provided by N, rather the subject of predication is assumed to do so). Consider the following sentence:

(25) *Der Doktor hält das für [den Bauch von vielen Schülern]*.
the doctor considers this for the stomach of many pupils.

In the example above the definite determiner has to be expletive w.r.t. LF, since the predicative NP is predicated of the demonstrative *das* which is therefore analyzed as binding the variable provided by N. On the other hand the nominal above cannot have the distributive inalienable interpretation. The lack of this interpretation indicates that in this configuration the definite determiner cannot be expletive w.r.t. domain D.[16] Without going into further detail I think that this piece of evidence can be taken as indirect support for the assumption that the (LF) 'binding approach' cannot be collapsed with the (domain D) 'instantiation-approach'. In addition, this fact provides also indirect evidence for assuming LF and domain D to be two distinct levels of representation.

2.2. Identification

Now we are at the point where we can start to explain what Identification (w.r.t. domain D) means. The various different properties will become clear when we start to analyze different sets of data. Consider again the representation for nominals used here:

(26) $[_{DP}$ the $[_{NP}$ man$]_X]_x$
NP: X = set of individuals satisfying the property denoted in N
DP: x = the discourse referent
$x \subseteq X$

Abstracting away from the contribution of the determiner[17] the only property on basis

[16] I do not have an explanation for why this should be the case. But notice that the reason for this can not be an inherent property of the NP under consideration since in the following sentence the same NP can have the distributive inalienable interpretation:
(i) *Der Doktor hat den Bauch von vielen Schülern untersucht*
the doctor has the stomach of many pupils examined
The lack of the distributive inalienable interpretation can also not be attributed to the fact that the 'subject NP' is singular. The same singular demonstrative can be subject of a plural NP as can be observed below:
(ii) *Der Doktor hält das für die Bäuche von vielen Schülern.*
the doctor considers this for the stomaches of many pupils
In general a neuter singular pronoun can be associated with a plural predicative NP (cf. section VII.3.2.2. for some further empirical evidence for this generalization.)

[17] They have to do with familiarity or novelty presuppositions on the one hand as well as uniqueness presuppositions on the other hand, besides various other contributions a determiner can make, but they are not relevant at the moment.

of which the actual discourse referent can be determined is the property denoted by N (i.e. in that case "being a man" or in our terms "belonging to the set of individuals who satisfy the property of being *a man*). To put it informally: especially for introducing a new discourse referent this is a very vague property to actually determine the relevant discourse referent. The universe of discourse (or the *contextual domain*) can contain more than one man. Picking out the particular man to be denoted might need some further specification, i.e. some further properties need to be ascribed in order to determine the actual referent. This is exactly what I think Identification really does here: it ascribes a property which suffices to pick the relevant referent out of the set of possible referents. This is also the reason for assuming that Identification is relevant at domain D - it is only w.r.t. a certain discourse that ascribing this property is relevant. For picking out an actual discourse referent domain D can induce the need for further licensing just in case the universe of discourse contains several possible individuals satisfying the property denoted by N. I will represent a DP which has to be identified as follows:

(27) $[_{DP}$ the $[_{NP}$ man$]_X]_x$
NP: X = set of individuals satisfying the property denoted in N
DP: x = ?
$x \subseteq X$

Although there is not enough content to pick out the discourse referent, there is one information we can rely on, namely that x has to be a subset of X. Now consider the representation of an identified nominal, i.e. let us provide another property which should suffice to identify the denotation of DP:

(28) *the man, who kissed Mary*

NP: X = set of individuals satisfying the property of being a man
DP: x = ?
$x \subseteq X$
CP: Y = set of individuals who have the property of "having kissed Mary"
$X \cap Y = x$

Ascribing another property to the intended discourse referent we can then create the intersection of X and Y. The property provided by the identifier should then suffice to actually denote the intended discourse referent. In case of definite DPs, which are associated with a uniqueness interpretation, the property denoted by the identifier should suffice to create a singleton set (by means of intersecting X and Y). Notice that the representation x = ? is also used in DRT. There it is referred to as an *incomplete*

condition, which are relevant for the analysis of anaphoric resolution. We will return to this issue and a comparison between Identification and binding in section III.5.

2.3. Identification and DRT

Let me briefly discuss the analysis of *directly referential nominals* within DRT. The DRT-treatment seems to involve a similar mechanism to what I think an identifier contributes to the interpretation of a nominal expression. DRT adopts the view taken for example in Kripke (1972) and Kaplan (1977) that elements like demonstratives, indexicals as well as proper names, are directly referential. This means that these elements contribute the same denotation in every possible world (thus they are not contingent). Within DRT this insight is implemented as to saying that there are *external anchors* (i.e. there is a function from discourse referents in the universe of a DRS to objects in the domain of the model). The individual that serves as the value for this function (the input of a discourse referent x) is called the *external anchor for x*. As noted in Asher (1993) the view of direct referentiality faces a problem for other assumptions within DRT. The problem is quite simple. If there are external anchors for discourse referents, why is it then necessary to have a condition in Con_K within the DRS? Why is it not enough to have the variable in U_K, i.e. what sort of information does the condition convey? To put it differently: why do we have to analyze names as predicates? This problem has to do with the assignment of truth conditions (i.e. the embedding into a model). Suppose that the above mentioned problem would be overcome along the following lines: a proper name could be analyzed as conveying some sort of metalinguistic information, namely that the discourse referent introduced by a name α stands for an individual with the name α. But this approach does faces a serious problem. Consider an obviously true statement like the one below:

(29) *Kripke is necessarily identical with Kripke.*

Assuming that there is some metalinguistic information (namely that *Kripke* conveys the information that there is an individual named *Kripke*) involved the sentence in (29) would falsely imply that the individual named *Kripke* necessarily has the property of *being named Kripke*. One cannot avoid this problem by simply eliminating the conditions for directly referential expressions. Obviously they do serve some purpose. This is also clear for complex demonstratives like *this boy*. According to Asher the relevant conditions serve the following purpose:

2.4. Identification as an (obligatory) licensing relation at domain D (empirical motivation).

There is a crucial argument for the assumption that the cases under consideration are to be analyzed as Identification at domain D. The argument is as follows. Consider again the examples which are taken to be the core cases of Identification:

(30) (a) *Peter hat es$_i$ geglaubt. ([daß Maria Bier trinkt]$_i$)*
Peter has it believed that Mary beer drinks

(b) *Peter hat daran$_i$ geglaubt. ([daß Maria Bier trinkt]$_i$)*
Peter has thereon believed that Mary beer drinks

(c) *Peter hat die Frau$_i$ getroffen. ([die gerne Bier trinkt]$_i$)*
Peter has the woman met, who likes beer drinking

(d) *Peter hat das Argument$_i$ vorgebracht. ([daß Maria gerne Bier trinkt]$_i$)*
Peter has the argument presented that Mary likes beer drinking

If one looks at the sentences in (30) in isolation it is not obvious that the (pro)nominals[19] require Identification. For pure syntactic well-formedness judgments all the examples are perfectly well-formed without the extraposed constituent (as indicated by the brackets). This is not surprising. Without embedding the sentences into a certain context the need for specifying the reference of a given NP is not obvious. The NP is inherently licensed (i.e. it has a theta- role and case is assigned, thus the syntactic purely syntactic licensing constraints are fulfilled).

However, establishing a relation like Identification in the theory of Grammar we would like to find elements which really show obligatoriness of Identification. Comparing Identification at domain D and binding (which might be viewed as Identification at LF) we find the following situation. The examples above show a similar behavior at domain D as pronominals do w.r.t. binding: pronominals can but need not be bound whereas the nominal elements above apparently can but need not be identified. However, as far as binding is concerned, there are elements which are lexically specified such that they need to be bound. This is true for lexical anaphors. Thus we might look for elements which are (lexically) specified for the necessity of being identified at domain D.

[19] Notice that I assume that *es* is a referential argument rather than an expletive (cf. chapter III.)

> *"They supply the agent with the means for connecting up the content of a message in which such expressions occur with his antecedent beliefs in the right way: the information conveyed by a proper name α for instance, allows the agent to integrate the information about the bearer of α with other information he has about the bearer."*
> (Asher 1993: 83)

And moreover for definite NPs Asher (1993) claims that:

> *"the semantic contribution (for the agent) of a definite NP, like a definite description or a complex demonstrative, is often simply to single out some individual for which a discourse referent has already been introduced in the processing of the discourse or which is otherwise available from the context, not to contribute to the set of property ascriptions attributed to the individual picked out."* (Asher 1993: 83)

Note that this is reminiscent of what we have said about the contribution of the identifier. Directly referential elements (as well as some other definites) are associated with conditions, but those conditions are of a different sort than the conditions discussed previously. For this reason Asher (1993) distinguishes between two sorts of conditions:

1) conditions that directly contribute to the truth conditions
2) conditions that are relevant to the interpretation by the agent[18]

In my view identifiers mainly convey conditions that are relevant for the interpretation by the agent. They denote properties that provide enough descriptive content to make the actual denotation available. They ascribe properties to an individual, but not (only) for assigning appropriate truth conditions, but also for picking out the intended referent. Notice that this is also the reason why I think that domain D (rather than LF) is the relevant level to capture the relation of Identification. Domain D as used here is independent of assigning truth conditions. Here I assume that LF is the relevant level. Again this implies that binding (to which we will return) is relevant at LF, since binding is necessary for evaluating the truth conditions.

[18] This obviously has a connection to Kripke's notion of Speaker's reference. The distinction between the two different conditions is also reminiscent of attributive vs. referential definite NPs (in the sense of Donellan 1966) (Cf. also Donellan 1978 for a discussion of the relation between attributive vs. referential properties of definites on the one hand and Speaker's Reference vs. Semantic Reference on the other hand).

As a matter of fact German has such elements: The relevant items are the complex determiners *derjenige, diejenige, dasjenige* which are special kinds of demonstrative determiners. The German examples below show the obligatoriness of Identification:

(31) (a) *Derjenige Mann kam auf das Fest *(der gerne Bier trinkt)*
det man came to the party who likes beer drinking
(b) *Peter hat dasjenige Argument vorgebracht *(daß Maria Bier trinkt)*
Peter has det argument presented that Mary beer drinking

These sentences are not wellformed without the identifier even when considered in isolation. The obligatoriness of the occurrence of another element seems to indicate that it is indeed the *cataphoric* element which needs some further licensing (i.e. Identification). Therefore I take these elements to crucially support the claim that Identification is a relation relevant for domain D.

According to Lehmann (1979) a similar phenomenon is found in Persian. Consider the following example and Lehmann's description:

(32) *Mardom-āzār-i-rā hekāyat kon-and ke sang-i bar sare-e sāleh-i zad*
people-teaser-IND-AKK tale make-3.PL [stone-IND at head-AT god-fearing-IND beat
'They tell about a 'people-teaser' who beat a God-fearer's head with a stone.'

" we saw that the head-noun bears the suffix -i. [. .] it is cataphorically similar to German 'derjenige' which also leads one to expect a relative clause."
(Lehmann 1979: 106, translation MW)

We can interpret this fact by saying that Persian marks obligatory Identification morphologically as indicated in the example above. In addition Persian has a special (demonstrative) determiner *ān* which seems to behave similar to German *derjenige*. Consider the following quotation of Lehmann (1979):

"Persian 'ān' works similar. Although it is a normal demonstrative. Since the head noun of a restrictive relative clause carries an indefiniteness-suffix, it follows from the combination of a definite and an indefinite in the same NP that the sentence is ungrammatical if there is no relative clause following."
(Lehmann 1979: 246, translation MW)

German does not only have determiners that induce the need for Identification. There are also other correlates that show this behavior:

(33) *Diese Schlußfolgerung ist insofern$_i$ nicht berechtigt *([als daß sie auf einer falschen Prämisse beruht]$_i$)*
this conclusion is insofar$_i$ not valid *([as that it on a wrong premise bases]$_i$)

(34) *Die Veranstaltung wurde dergestalt$_i$ konzipiert, *([daß man ein Ende nicht absehen konnte]$_i$)*
the event was this-way$_i$ planned *([that one an ending not foresee could]$_i$)

I think that this phenomenon (of obligatory cataphoricity) provides crucial evidence for assuming Identification as the relevant relation between the nominal element and the sentential constituent.

Notice that there is another approach to be found in the literature for explaining the relation between the nominal element and the (extraposed) clause, which cannot really capture the facts we are dealing with here. For the moment I will just summarize the main points of this analysis and its shortcomings. In Appendix 1, I will compare the two approaches more carefully. It is assumed for example in Culicover & Rochemont (1990) (henceforth C&R) (following Guerón 1980) that the relation between the nominal element and the extraposed constituent is an instance of *complementation*. This relation is syntactically relevant (either at S-structure (C&R) or at LF (Guerón & May 1984). Under this approach the examples in (33-34) would be an instance of obligatory complementation. The first argument against such a view has to do with the remarks on licensing discussed in chapter I. Remember that I assume that it is never the case that an element X has to license another element Y. It is always the case that an element needs to be licensed. However, in the example above one would have to claim that the nominal element has to obligatorily license the extraposed clause as its complement. This already leads to the second problem. Being an obligatory complement is reminiscent of 'being selected', i.e. being an argument. In that case the determiner would have to be analyzed as a two-place-predicate. However this assumption would undermine the introduction of the complement-relation as a primitive relation. C&R argue that

"the complement to a maximal projection [..] is a phrase that holds an adjunct or argument relation to the head of the phrase it takes as antecedent".
(C&R 1990: 26, Fn.9)

This means that it is the fact that it is NOT only (selected) arguments that are relevant for the Complement Principle. The third argument against assuming complementation as the relevant relation is the following. The paradigm in (30-32) shows that Identification in general is just necessary if the content of the nominal is not specified

enough to pick out the actual discourse referent. Thus we have seen that Identification is in general syntactically optional. However, Identification might turn obligatory (depending on discourse properties or on the choice of the determiner). Complementation (as a purely syntactic relation) should not be sensitive to discourse phenomena.[20]

The last point becomes even clearer on basis of the following paradigm. An equivalent device to using the complex determiner is (contrastive) stress on a definite determiner or on the pronominal element respectively. The following examples exemplify this behavior:

(35) Stress on the definite determiner
  (a) *DER Mann wollte auf ein Fest gehen *(der gerne Bier trinkt)*
  THE man wanted to a party go who likes beer drinking
  (b) *Peter hat DAS Argument vorgebracht *(daß Maria Bier trinkt)*
  Peter has THE argument presented that Mary beer drinking

(36) Stress on the pronoun
  *Peter hat {DAran}, geglaubt, *({daß Maria gerne Bier trinkt},)*
  Peter has THEREon believed that Mary likes beer drinking

[20] This statement might be too general (as pointed out to me by E.Williams). Whether a complement needs to be specified can be discourse related. However, I think that there is still something else going on. Consider the following example:
(i) *Ich habe getrunken.*
I have already drunk
Sentence (i) involving the transitive Verb *trinken* is wellformed without the direct object. There are two possibilities when this sentence is wellformed. Either *drinking* refers to being an alcoholic, thus it is necessarily interpreted as *drinking alcohol*. This is the interpretation one gets if the sentence is uttered in isolation. In that case it cannot mean that there is something which you normally do not drink. However, embedded in a discourse, where it is clear what was drunk (e.g. *poison*, something you do not drink under normal circumstances) then the sentence above is also well-formed, indicating that complements can be rendered optional due to discourse properties. Still I think that this examples differs to the ones where I think that Identification is at play. The first difference is that all examples involving Identification are perfectly well-formed without the identifier when they are considered in isolation. They do not even receive a special interpretation. It is just the case that a certain discourse can render Identification obligatory (or superfluous). On the other hand there are many examples where the occurrence of the (selected) complement is necessary, even if the denotation of the complement is clear from the context.
(ii) S1: *Ich habe das Buch der Maria gegeben*
I have the book to Mary given
S2: *Ich habe *(das Buch) dem Hans gegeben*
I have *(the book) to John given
We can conclude that complementation in general can be sensitive to discourse phenomena, but the facts seem to be different to the cases we are interested here.

Again making use of a complement relation would be problematic, since focus should not in general induce the need for the obligatoriness of a purely syntactic relation like complementation. We will see in section IV.7, how the analysis presented here derives these facts concerning (contrastive) stress and the obligatoriness of Identification. It has to do with the fact that stress on the determiner explicitly indicates that there has to be a distinguishing property, which is subsequently provided by the identifier.

3. The identifyee introduces a discourse referent

Let me now address the second assumption namely that *the identifyee introduces a discourse referent*. Therefore it is subject to the Novelty Condition (resulting in a precedence Constraint). The Novelty Condition originally formulated in Heim (1982) is given below

(37) Extended Novelty-Familiarity Condition:
For ϕ to be felicitous w.r.t F it is required for every NP_i in ϕ that:
(i) if NP_i is [-definite], then $i \notin$ of Dom(F)
(ii) if NP_i is [+definite], then
  a) $i \in$ Dom(F), and
  b) if NP_i is a formula, F entails NP_i (Heim 1982: 369f.)

Informally spoken: if the identifyee is linearly preceded by an element which bears the same index it violates the Novelty Condition. There are two necessary assumptions for this proposal to work. However, I think that both are independently necessary. The first assumption is that sentential elements are associated with a discourse referent. This is necessary because the Novelty Condition (resulting in the precedence constraint) shall explain the obligatory rightness of the sentential element. Therefore it has to be coindexed with the nominal element. The second necessary prerequisite is to assume that the Novelty-Familiarity-Condition is evaluated linearly in a step by step manner and not in an 'all-at-once' manner. Since both assumptions are not so obvious within standard (DRT)-treatments I will briefly discuss and justify them.

3.1. Sentential elements are associated with a discourse referent

At first glance it is not a straightforward assumption, that CPs are associated with a discourse referent. Notice that Heim (1982) explicitly formulates the Novelty-Familiarity Condition as to hold only for NPs. Moreover V&Z assume that it is just

nominal elements which are represented at domain D (i.e. L-structure).

In this section I want to show that this cannot be the whole story about entities that are associated with discourse referents. This claim is very crucial for several aspects of the present analysis. I will argue that in addition to NPs also propositions as well as other abstract entities (e.g. properties, amounts, results) are associated with a discourse referent.

I think that a crucial problem with the version of the Novelty-Familiarty Condition (formulated to hold just for NPs as in (37)) is the following implication it might suggest: In a way it presupposes that different syntactic categories map to different ontological entities in a one to one relation. However, this is in general not the case (as we have already discussed in section II.2.). The second problem has to do with a prerequisite for the postulation of domain D. Referents at domain D are not to be equated with the notion of reference in philosophy. In talking about entities relevant at domain D we are not talking about "concrete" or "existing" entities in the real world. Domain D is not to be equated with the real discourse[21] where physical objects are present, objects we can touch, point to and of course talk about. We are also able to talk about *abstract entities*, about properties, concepts, events, states of affairs, beliefs etc.[22]

There is a simple argument in favor of the claim that abstract entities like propositions are associated with a discourse referent und thus bear an index relevant for domain D. It cannot be denied that a pronoun can take a whole proposition - and thus a CP - as its antecedent. This is shown in the example below:

(38) *[Maria hat gestern ein Bier getrunken]$_i$ aber Hans hat [dastes]$_i$ nicht geglaubt.*
'[Mary has drunk beer yesterday], but John has not believed [that/it],'

Some more examples exemplifying this phenomenon are given below. They are taken from Asher (1993) and provide evidence that all sorts of different abstract (or semi-concrete) entities can be used as antecedents for anaphoric linking (as a source for these data he cites *The Wall Street Journal*):

[21] Notice that domain D is also not to be equated with models, or possible worlds. It is an abstract level of representation only relevant for Grammar.

[22] For an elaborate discussion of reference to abstract objects in discourse I refer the reader to Asher (1993). There abstract entities like *eventualities* and *propositions* (among others) are discussed and argued to be associated with a discourse referent.

(39) (a) *Be careful what you wish...because wishes sometimes come true. **That's** what the Semiconductor Industry Association, which represents U.S. manufacturers, has been learning.*

(b) *Well, a clerk told us, we'd need to hire a lawyer to make a petition--but it probably wouldn't be worth the effort...*

(c) *Being spanked by Milton Friedman is one of life's most humiliating experiences (see alongside), so we feel compelled to fess up about when we'll proclaim the "right" price for ice-cream cones. We faithfully promise to do **that** on the day Milton Friedman proclaims the "right" percentage for annual growth in the ice-cream cone supply.*

(d) *As part of corporate streamlining programs, many companies are extending early-retirement packages to legions of senior managers. They see it as one relatively painless way to pare management ranks...*

(e) *But "Misery," which runs its mostly gripping if unattractive course in one little bedroom with only two real characters, is selling like mad. It just goes to show you that some people don't know when they've been insulted.*
(Asher 1993: 3)

The (bold) pronouns (in 39) anaphorically refer back to propositions (39a), a property or type of action in (39b-c), a type of event in (39d), and a fact in (39e).

Moreover it is generally agreed on that pronouns cannot introduce a new discourse referent. If however the CPs above would not bear an index, then the pronoun (which should bear an index as any other NP) would introduce this index. In addition it is not clear how the anaphoric link between the proposition and the pronoun should be established if the proposition would not count as its antecedent, i.e. if the CP would not be an entity relevant for domain D. Therefore I think that it is safe to conclude that CPs are associated with a discourse referent.[23] Asher (1993) gives a variety of other facts leading to this conclusion, which I will not discuss here.

[23] The assumption of CPs being associated with a discourse referent would open another field of investigation. One could suspect that the Novelty-Familiarity Condition also applies to those entities. If so then there might be different instantiations of sentential elements, depending on whether they contain old or new information. Notice however that for the purpose of the present analysis we can abstract away from this possibility. In the examples we are considering here it is the correlated nominal element which is assumed to obey the Novelty Condition (and not the sentential identifier).

There is one more issue to address. The following paradigms are independent support for the claim that there is a step by step calculation concerning the Novelty-Familiarity condition which heavily relies on S-structure. There is evidence that the possibility to reconstruct (i.e. to calculate the element within its base-position) is excluded: Novelty-Familiarity violations cannot be saved via reconstruction (as opposed to binding effects which sometimes can be calculated in the reconstructed position).

A'-movement in German and English show reconstruction effects. We find the possibility to interpret a given constituent in its overt (S-structure) position or in its D-structure position:

(42) (a) *Welches Bild von sich$_i$ hat Hans$_i$ gemalt*
Which picture of himself has John painted
(b) *Welches Bild von sich$_i$ glaubt Hans$_i$ daß Peter$_j$ gemalt hat?*
Which picture of himself believes John that Peter painted has
(43) (a) *Which pictures of himself$_i$ does Jane think that John$_i$ will sell*
(b) *Which pictures of himself$_{i/j}$ does John$_i$ think that Bill$_j$ will sell*

The anaphor in the wh-moved constituent is not in the c-command domain of a possible antecedent. Nevertheless the sentence is wellformed. Thus it is assumed that the wh-moved element can be 'reconstructed' at LF - where binding theory is assumed to apply. However as the following examples show this is not so for the Novelty Condition:

(44) (a) *?Welches Buch eines Lehrers$_i$ hat Hans demdiesem Lehrer$_i$ geklaut?*
Which book a-GEN teacher has John the teacher stolen?
(b) **Welches Buch desdieses Lehrers$_i$ hat Hans einem Lehrer$_i$ geklaut?*
Which book the/this-GEN teacher has John a teacher stolen

Although sentence (44a) is not perfectly wellformed there is a clear contrast: the sentence in (44b) is impossible under the coreferent reading but the sentence in (44a) is quite acceptable. So we have independent support for the claim that first the S-structure position is relevant for the evaluation of the Novelty-Familiarity-Condition and secondly, that the Novelty Condition is evaluated in a step-by-step manner. If it would be evaluated only on basis of the whole structure (i.e. the entire sentence), then the contrast in (44) can not be explained. The sentence is only well-formed if the indefinite DP linearly precedes the coreferent definite DP. Although A'-movement in general shows reconstruction effects, the above sentences show that violations of the Novelty Condition cannot be saved via reconstruction. In the next chapter we will see another piece of evidence for this claim having to do with backward-pronominalization.

3.2. Linear step-by-step evaluation of the Novelty Condition

The second necessary assumption is to say that the evaluation of the Novelty Condition is done linearly from left to right. This is however an assumption implicitly made in Heim (1982). It is not after perceiving the whole structure but after each relevant occurrence of an indexed NP that the condition is checked. Consider for example the following discourse:

(40) *A woman was bitten by a dog. She hit him.*

There are four relevant 'formulas ψ' each containing one or more files involved: *[a woman] [a dog] [e was bitten by e]* and *[she hit him]*. Heim assumes successive application of the rule to calculate the Novelty-Familiarity-Condition. Calculation proceeds as follows:

(41) $F+\psi^1$
$(F+\psi^1)+\psi^2$
$((F+\psi^1)+\psi^2)+\psi^3$
$(((F+\psi^1)+\psi^2)+\psi^3)+\psi^4 = F + (X))$

"At each of these four steps, the Novelty-Familiarity-Condition has to be met before the calculation can be carried out" (Heim 1982, p. 322).

I do not want to go into a detailed discussion of her analysis. The above remarks should suffice to see the implication that calculation precedes linearly in a step by step manner. I differ from Heim's (1982) approach in assuming that it is the S-structure position which is crucial for evaluating the Novelty Condition whereas Heim assumes LF to be the relevant input.[24]

For the examples we are interested here this simply means that in order for a constituent introducing a discourse referent x to be felicitous x cannot be already present in domain D at this point of the calculation. However if the sentential identifier occurs in a position preceding the nominal this is not the case anymore.

[24] Notice that the approach I am advocating here is also supported psycholinguistically by Chater, Pickering & Milward (1994). They argue that interpretation proceeds in an *incremental* way. There is evidence that once a constituent is uttered (i.e. heard) the process of understanding starts and is not delayed until the whole sentence is perceived.

III. IDENTIFICATION OF PRONOMINALS

We are now in the position to see how the present proposal applies to the data under consideration. Let us start with sentences involving a pronominal element in the matrix. Some examples are given below:

(1) (a) *Peter hat es geglaubt, daß Maria gerne Bier trinkt.*
Peter has it believed that Mary likes beer drinking
'Peter believed that Mary likes beer drinking'

(b) *Klara hat es bedauert, daß Egon verschwunden ist.*
Klara has it regretted that Egon disappeared is
'Klara has regretted that Egon has disappeared.'

(c) *Egon hat es verstanden, den Hans zu ärgern.*
Egon has it understood, the John to annoy
'Egon understood how to annoy John.'

(d) *Egon hat es schon erfragt, ob Klara Bier trinken darf.*
Egon has it already asked, whether Klara beer drink may
'Egon has already asked, whether Klara is allowed to drink beer.'

(e) *Maria hat sich darüber aufgeregt, daß Hans kein Bier trinkt.*
Mary has REFL thereon complained that John no beer drinks
'Mary complained about John not drinking beer'

(f) *Klara hat sich danach erkundigt, wie der Berg zu besteigen ist.*
Klara has REFL thereafter asked, how the mountain to climb is
'Klara asked how to climb the mountain'

In all of these sentences we find an extraposed clause which is correlated with a pronominal element in the matrix clause. There are two main approaches to analyze the relation between the two elements. One possibility is to say that the pronominal in the matrix is an expletive element and the extraposed clause is really the argument of the verb (cf. Postal & Pullum 1988 for English). The other possibility is to say that the pronominal is really the argument. This second approach allows for two possibilities to analyze the status of the correlated sentential element: either it is treated as a complement of the pronominal itself (cf. Müller 1993) - an analysis I will reject. The other possibility is to treat the clause as an apposition (cf. Cardinalletti (1990) as well as Hoekstra 1983 and Bennis 1986 for Dutch). I will adopt the latter proposal, i.e. that syntactically the clause is licensed as an apposition but that there is more to this relation: at domain D this appositional relation (which is not an inherent licensing possibility) is licensed as an instance of Identification. In assuming this we will see in section III.3. that in a specific way the *expletive* character of the element does play a

role. It is simply the case that a given element can act as an expletive w.r.t. one level of representation (in the sense of V&Z, cf. section II.2.1.1.) but it is licensed at another level of representation, where it can therefore not be analyzed as an expletive. We will see in section III.3. how this assumption explains the apparent problem concerning *optionality*.

Let us now look at some properties of the sentences under consideration. First a few words about the terminology are in order. Since the pronominal elements within the matrix are correlated with a clausal element it is not fully appropriate to talk about *pro-nominals*. Moreover in the sentences in (1) the correlated element is categorically a PP (but also coreferent with a clausal element). Within traditional Grammar these elements are referred to as *pronominal adverbs*. In more recent theories they have been labelled ProPPs[1] (cf. Altmann 1981, Wiltschko 1993). However, as has been argued in various places (Riemsdijk 1978, Wiltschko 1993, Müller 1993) these ProPPs involve a (syntactically transparent) pronominal, i.e. *da*. As for the problem that they are correlated with a sentential element there is one interesting point to make. Within several (philosophical and linguistic) approaches the following elements are referred to as *sentential nominals* (cf. Asher 1993):

(2) *of-ing* gerunds, derived nominals, *that*-clauses, *for*-infinitivals, naked infinitives
Poss-ing gerunds, *Acc-ing* gerunds.

This terminology is intended to summarize elements denoting abstract objects - besides sentential nominals there are also predicative nominals which show these properties. We have already discussed the fact that there is no one to one mapping between syntactic categories and ontological categories. All the instances of clausal elements we are interested in (finite and infinitival clauses) can be anaphorically linked to a (pro)nominal. I will therefore continue to use the term 'pronominal' for these instances of correlates as opposed to other full lexical NP correlates.[2] For the categorical nature of the correlate I will assume that this is a matter of c-selection (by the verb).

What interests us now are the properties of Identification and whether these properties are attested in the examples above. One of these properties we have already seen is that the occurrence of the identifier is optional (for pure syntactic

[1] This term has initially been introduced in Kayne (1975) for the French elements *en* and *y*.

[2] I will abstract away from the differences of the sentential elements i.e. whether it is a finite clause or an infinitival or what sort of complementizer it takes respectively. (see among others Hegarty 1990 for an analysis).

wellformedness judgements). The same is true for Identification of pronominals as can be observed below. Here we find the same sentences as in (1) with the only exception that the (extraposed) clause is missing:

(3) (a) *Peter hat es geglaubt.*
'Peter believed it.'
(b) *Klara hat es bedauert*
'Klara regrets it.'
(c) *Egon hat es verstanden*
'Egon understood it.'
(d) *Egon hat es schon erfragt*
'Egon has asked for it.'
(e) *Maria hat sich darüber aufgeregt*
Mary complained about it.'
(f) *Klara hat sich danach erkundigt*
'Klara has asked for it.'

Syntactically all these sentences are perfectly wellformed. However, for the interpretation it seems clear that in all of the above cases the content of the pronoun needs to be clear from the context. This is not a surprising fact since pronouns need to be licensed by a previous discourse referent. But now we seem to have an apparent problem. On the one hand it is a crucial assumption within the proposal suggested here that the identifyee introduces a discourse referent but on the other hand it is standardly assumed that pronominals cannot introduce a discourse referent. However, there is a way to maintain both claims, which in fact our analysis forces us to assume. This will in effect turn out to be a rather desirable claim since some otherwise unexpected properties of the cases under consideration are derived.

1. Pronouns introducing discourse referents

Let us use look at some of the examples in more detail in order to exemplify the analysis. In accordance with what we have said about the representation of nominals I will assume that pronominals are also associated with two indices (as indicated below). For the syntactic representation this assumption implies that the structure of pronominals is as follows:

(4) $[_{DP}\ D^{\circ}[_{NP}\ N^{\circ}]_X]_x$

It has to be assumed that pronominals project like other nominal elements: an NP as well as a DP. Furthermore I will assume that the pronominal is basegenerated in N° and then moved to D° via head to head movement. (Or maybe in case of 'strong' pronouns the whole NP might move to SpecDP).[3] This differs from analyses of pronominals which claim that they are to be analyzed as intransitive D°s. However assuming some version of 'Extended Projections' (cf. Grimshaw 1990) I think that this view is forced anyway since functional projections are only licensed as projections of lexical categories. Since it is not in the scope of the present discussion I will not go into more detail. We can now assume the following representation for the sentences under consideration:

(5) (a) *Peter hat [es]$_{x_i}$ geglaubt, [daß Maria Bier trinkt]$_i$*
Peter has it believed that Mary beer drinks
(b) *Peter hat [daran]$_{x_i}$ geglaubt, [daß Maria Bier trinkt]$_i$*
Peter has thereon believed that Mary beer drinks

Now we are faced with the problem to explain what the denotation of NP (X) in case of pronominals corresponds to. For full lexical NPs I have argued that X denotes the set of individuals satisfying the property denoted by N. However, obviously a pronoun does not have any descriptive content (except for its phi features), i.e. there is no N denoting a certain set of individuals. I will assume for X of pronominals to denote *the set of possible antecedents*. Thus, the standard claim that pronouns cannot introduce a discourse referent is reinterpreted as to saying that it is associated with an index (X) which has to denote a set of "possible (already established) discourse referents". If one assumes that the pronoun moves from N° to D° then it could be argued that the trace of the pronoun leaves the relevant restriction which come with their phi-features. The difference between lexical NPs and pronominals is then that the former restrict the set of possible discourse referents by means of N and the latter do so by means of their phi-features (i.e. gender and number).[4] On the one hand this still implies that a discourse referent must already be established since a set of possible antecedents must be available. However, the actual discourse referent can be contained in a whole set of possible discourse referents. In some cases we will find that if the discourse referent

[3] One could also claim that the pronominal is basegenerated in D° and selects for an empty NP. For the present discussion nothing would change. Cf. Koopman (1993) for an elaborate analysis of pronominals.

[4] Notice that this analysis can also account for the fact that in many cases the occurrence of a pronoun is ambiguous. Thus if more (salient enough) discourse referents are available which match with the pronoun's phi-features then all of these discourse referents are in principle possible to count as the denotation of the pronoun.

x is contained within the set X it is not accessible for the pronoun to being picked up. The actual referent x still has to be identified.[5] In addition there is the restriction that x has to be a subset of X. In other words x has to be re-introduced in domain D, although x is already established it is not immediately accessible to anaphoric linkage. As we will see in case of definite DPs (and many pronominals are definite nominals) the contribution of x is to provide a *uniqueness* interpretation. This means that the occurrence of the pronominal indicates that there has to be a unique discourse referent. The identifier provides enough descriptive content to pick out the intended referent, i.e. it provides discourse availability. A sentence like (5a) can be represented as follows (with some abbreviations which are not relevant at the moment):

(6) Pron: X = the set of possible discourse referents
x = ?
$x \subseteq X$
CP: x = 'that Mary drinks beer'

There are some predictions inherent to this claim which I will discuss in the subsequent sections.

2. Presuppositions induced by $x \subseteq X$

If there has to be a set of possible (already established) discourse referents which the pronoun refers to it is predicted that the content of the pronoun is presupposed. Here we are dealing with the notion of presupposition in the sense of Karttunen (1974), i.e. that an utterance U presupposes a proposition P if that utterance is acceptable only at a point in discourse where that proposition is in the set of propositions that the parties to the discourse take as established. Notice that this assumption is in the spirit of Williams (1994). We are dealing with a presupposition concerning the existence of an antecedent. Williams (1994b) equals (this sort of) presuppositions with anaphoricity.

There are various ways to test for presuppositions which I will now present. They all show that the prediction is borne out.[6] The first test involves question-answer pairs. Consider the following examples:

[5] cf. section III.5. for a more detailed discussion of this claim.

[6] Cf. also Iatridou (1993) for English that the presence of the pronominal *it* which is correlated with a sentential element implies familiarity.

(7) Q: *Was hast du geglaubt?*
What did you believe?
A1: **Ich habe es geglaubt, daß Hans Bier trinkt*[7]
I have it believed, that John beer drinks
A2: **Ich habe daran geglaubt, daß Hans Bier trinkt*
I have thereon believed that John beer drinks
A3: *Ich habe geglaubt, daß Hans Bier trinkt*
I have believed that John beer drinks

The question in (7) is about the content of the addressee's belief. Thus it implies that this content is not presupposed, i.e. that it has not been established in the discourse yet. Therefore an appropriate answer will introduce this content of belief as a new discourse referent. As we expect the only appropriate answer for a question like this is the one without the pronominal in the matrix, since the pronominal is just appropriate if there is a set of available (already established) discourse referents.[8] Consider now an example where the question makes a discourse referent available:

(8) Q1: *Hast du geglaubt, daß Maria gerne Bier trinkt?*
'Did you believe that Mary likes to drink beer?'
Q2: *Wer hat geglaubt, daß Maria gerne Bier trinkt?*
'Who believed, that Mary likes to drink beer?'
A1: *(Ja), ich habe es geglaubt, daß Maria gerne Bier trinkt*
'Yes I believed it that Mary likes to drink beer'
A2: *(Ja), ich habe geglaubt, daß Maria gerne Bier trinkt*
'Yes, I believed that Mary likes to drink beer'

[7] There is one reading where the presence of the pronominal in the answer is appropriate. This is the case if there are several possible propositions available and the question is about 'which of these possible propositions was the one the addressee believed'. This corresponds to a d-linked question (in the sense of Pesetsky 1987). Therefore it is not a problem for the analysis presented here.
The d-linked reading can however be excluded with "aggressively non-d-linked" wh-phrases as in the question below:
(i) *Was zum Teufel hast du geglaubt?*
What the hell did you believe?

[8] Notice that the relevant notion here is really familiarity rather than factivity. This can be seen from the fact that the same kind of presupposition is induced with factive verbs:
(i) Q: *Was hast du bedauert?*
What did you regret?
A1: **Ich habe es bedauert, daß Hans Bier trinkt*
I have it regretted, that John beer drinks
A2: *Ich habe bedauert, daß Hans Bier trinkt*

In case the content of the belief is already established it is legitimate to have the pronominal in the matrix. Notice that this does not contradict our claim that *es* still (re-)introduces a discourse referent. In general within an answer one can repeat an NP which has been 'introduced' in the question. This can be observed in the following question answer pair:

(9) Q: *Hast du ein Buch verbrannt*
'Did you burn a book?'
A: *Ja, ich habe ein Buch verbrannt.*
'Yes I have burnt a book.'

Although the NP *ein Buch* is already mentioned in the question it is nevertheless repeated as an indefinite NP - apparently contradicting the Novelty Condition. This suggests that discourse referents cannot be introduced in the context of a question. I have nothing more to say about this fact. It just shows that there is no problem posed by (8) for our analysis concerning the Novelty of the pronominal.

The second test which shows the same phenomenon involves focussing a constituent within the identifier. Since Focus is related to new information on the one hand but on the other hand the content of the identifier is presupposed we predict that the identifier cannot contain a focussed constituent. This is again a correct prediction:

(10) (a) **Peter hat es gelaubt, daß MARIA Bier trinkt*
Peter has (it) believed, that MARY beer drinks
(b) *Peter hat gelaubt, daß MARIA Bier trinkt*
Peter has believed that Mary beer drinks

(10) shows that an extraposed clause associated with a pronominal cannot contain a focussed phrase. On the other hand the sentence is wellformed in case no pronominal is present in the matrix. Notice that there is a reading available under which the sentence in (10a) is wellformed. Under this reading the focus in the extraposed clause is interpreted as an 'Echo-focus'. Thus the proposition already established must be a proposition where Mary is contrastively focussed.[9]

[9] A similar phenomenon is discussed in Williams (1994):
"...the presence of a tonal marking in the anaphor gives the sentence a very specialized "metalinguistic" use - it can only be used when previous discourse provides an appropriate antecedent, one with the same tonal marking" (Williams 1994: 31).
He gives the following discourse exemplifying this phenomenon:
(i =92) A: *Mary likes the short policeman, but the uncooperative prisoner likes the TALL policemen.*
B: *no, the COOPERATIVE prisoner likes the TALL policemen.*
And furthermore Williams states that:
"B has, as it were, picked up as much as the surface of A's sentence as possible, including stress contour; this is why it has an "echo" quality." (Williams 1994: 32)

Finally, on basis of the same line of reasoning we also predict that it is impossible to construe a focus particle (like *sogar*) with an extraposed clause if the latter is coindexed with a pronominal in the matrix[10] . Again this is a correct prediction of the present analysis: the focus particle would indicate new information but the content of the pronominal must already be established in the discourse:

(11) (a) **Peter hat es sogar geglaubt, daß Maria Bier trinkt*
Peter has it even believed that Mary beer drinks
(b) *Peter hat sogar geglaubt, daß Maria Bier trinkt*
Peter has even believed that Mary beer drinks

Let me now address some other predictions the present analysis makes. They have to do with the representation of pronominals (assumed here) combined with the assumption that domain D is the level where Identification is relevant. The following discussion will provide an interesting result concerning the issue of optionality.

3. Optionality

So far we have not said anything about extraposed constituents which are not related to a correlate in the matrix. In that case the extraposed clause cannot be analyzed as an instance of Identification - at least not in the same manner as in the examples we have been considering so far. For completeness let me briefly say how I think these instances are licensed. If we look at purely syntactic well-formedness conditions the presence of the pronominal seems to be optional in many cases. Consider the following examples:

(12) (a) *Peter hat es geglaubt, daß Maria Bier trinkt*
Peter has it believed that Mary beer drinks
(b) *Peter hat geglaubt, daß Maria Bier trinkt.*
Peter has believed that Mary beer drinks
(c) *Peter hat es geglaubt.*
Peter has it believed
(d) **Peter hat geglaubt.*

[10] This phenomenon was pointed out to me by Hubert Haider, p.c.

The paradigm above shows that either of the two elements (the nominal and the sentential constituent) must be present syntactically. If neither is present the sentence is ungrammatical (12d). However syntactically it seems optional whether the pronominal or the sentential constituent or both of them are present. I will assume that extraposed constituents which are not correlated with a (pro-)nominal element in the matrix are also licensed as identifiers. However here we find an instance of 'syntactic' Identification. This can be interpreted as an instance of theta-Identification. It is also somehow in the spirit of Williams (1994), who assumes that theta-assignment can be interpreted as an instance of binding. As we will discuss in section III.5. binding and Identification are similar notions. Assuming this, the different properties of extraposition with- or without a correlate follow straightforwardly. The obligatory occurrence of the clause follows since it is necessary for syntactic reasons (i.e. theta-assignment). As we will see in section VI.4. there is no precedence constraint in these cases since PCI only applies to overt elements (i.e. elements which introduce a discourse referent). Finally θ-Identification and Identification at domain D differ in their Locality Domain. According to Higginbotham (1985) θ-Identification takes place under government. Therefore I will assume that the extraposed clause is adjoined to VP.[11] For the moment these remarks should suffice in order to discuss the issue of optionality more carefully.

The problem of optionality is as follows. As we have said the occurrence of the pronominal seems to be optional for purely syntactic wellformedness judgements. In general optionality is a problem in recent developments within syntactic theory. For example Chomsky (1992) argues that movement is never optional. Since movement and insertion are both analyzed as instances of 'generalized transformation' we can extend this proposal for 'insertion' and thus for the examples in (12). Obviously this relates to the Principle of Full Interpretation which is interpreted as claiming that every element needs to be licensed. On the other hand it is also a very traditional claim that any overt (syntactic) difference should imply a difference in meaning. This traditional claim is for example used as an argument for the relevance of *blocking* in syntax in Williams (1994b). There he argues that

> *"rather than taking 'sameness of meaning' to be a condition delimiting the application of the blocking principle, we could turn things around, and make 'sameness of meaning' the heart of the blocking principle: 'if two forms are different, then there must be a difference in meaning (so find it)".* (Williams 1994: 4)

[11] cf. Wiltschko (1993, 1994) for a more detailed discussion

Difference in meaning can be of semantic or of pragmatic character. What does this mean w.r.t. to the data we are interested in here? As we have seen the optionality of the presence of the pronoun is just an apparent one, i.e. it just concerns syntactic wellformedness judgements. Once both variants of the sentences above are embedded into a certain environment, which forces or forbids a certain interpretation, the problem of optionality disappears. Therefore the occurrence of *es* in (12) is really not arbitrary. It does encode some difference in meaning which is pragmatic in nature, i.e. the sentences have different felicitous conditions. Similar conclusions are detected for a variety of phenomena within syntactic theory. For example it has been a longstanding claim that German has free word order (referred to as *scrambling* within the generative theory). This phenomenon is only apparently a completely 'free-word-order'. Again for purely syntactic reasons several linearization possibilities within the German *Mittelfeld* are to be observed. However, if embedded in certain contexts the apparent optionality again disappears. Different word orders correspond to different 'meanings' having again to do with presuppositions. Much of recent work within syntactic theory is devoted to this attempt to explain apparent optionality (cf. among others Diesing 1992 for explaining apparent 'scrambling' phenomena along these lines).

There are two important remarks w.r.t. to this general development within syntactic theory. First (as we have already discussed) a strict autonomy version of syntax seems to be problematic.[12] Although it is true that syntactically there seems to be optionality at play in several cases - thus the syntax module rules in a given phrase marker which is ruled out (by some other wellformedness condition) once the sentence is embedded in a certain context. These instances, where a syntactic wellformed sentence is ruled out in certain contexts, need an explanation within the theory of Grammar. It is thus the problem of encoding these phenomena in a framework which postulates the autonomy of syntax. The problem is that syntax in general has to generate the structure but in one way or another the representation must be able to derive the well- or ill-formedness of the sentence embedded in a certain discourse.

3.1. Taking discourse into account: Functionalism vs. Formalism

At this point I want to discuss Functional Grammar (as for example developed by Dik 1978) as well as functional explanations in general in some more detail. I will do so

[12] Notice that in Chomsky (1992, 1994, 1995) a strict version of the autonomy hypothesis seems to be given up, i.e. there it is claimed that constraints for representations are argued to be constraints imposed by the *Interface Levels*.

since the analysis suggested so far has shown that discourse phenomena have to be taken into consideration. Doing so might seem contradictory to standard generativist assumptions where it has been explicitly denied that facts related to discourse are relevant for syntactic analyses (the *Autonomy Hypothesis*). Functional Grammar on the other hand is an approach where discourse phenomena are taken to be the most crucial factors for explaining language. Thus I will briefly introduce the main claims within Functional Grammar. Then I will show why such an approach does not work and finally I will briefly discuss an analysis within the generative approach in order to account for the data under consideration (in a way which does not contradict the Autonomy Hypothesis).

3.1.1. Functional Grammar

There are two sides of the coin, i.e. there are two 'functional' properties of Functional Grammar. One is to be viewed within a general theory of Functionalism. This opposes Functional Grammar to the *formal* paradigm. The relevant properties concerning this part of the coin are listed below:

- Language is taken to be an instrument of social interaction
- The primary function of a language is communication
- The psychological correlate of a language is communicative competence: the ability to carry on social interaction by means of language
- The study of the language system must from the very start take place within the framework of the system of language use
- The description of linguistic expressions must provide points of contact for the description of their functioning in given settings
- The child discovers the system underlying language and language use, aided by an extensive and highly structured input of linguistic data presented in natural settings
- Language universals are to be explained in terms of the constraints inherent in (i) the goals of communication, (ii) the biological and the psychological constitution of language users, (iii) the settings in which language is used
- Pragmatics is the all-encompassing framework within which semantics and syntax must be studied: semantics is subservient to pragmatics and syntax to semantics; the priorities run from pragmatics via semantics to syntax. (cf. Dik 1978: 5)

These properties are related to Functionalism in general by means of trying to explain a given object (language) through ascribing it a function (communication/social interaction) which should suffice to explain the object itself. What interests us here are mainly two points (differentiating this approach from the generative framework):

a) that Functional Grammar tries to explain language by means of the main function they ascribe to it (i.e. language as an instrument of social interaction. Explanations start from looking at language use.) and
b) because of this point of view it is assumed that pragmatics has priority over semantics and syntax (the latter being supervient).

The second side of the coin concerns some properties inherent to Functional Grammar, namely that the rules are formulated in terms of functional (and categorical) properties of constituents. Three levels of functions are assumed:

(i) Semantic Functions: AGENT, GOAL, RECIPIENT, etc.
(ii) Syntactic Functions: Subject and Object
(iii) Pragmatic Functions: Theme and Tail, Topic and Focus

These functions are assigned to a given representation (which is derived via predicate formation). They are considered to be primitives. The following representation is a possible output of predicate formation and subsequent semantic, syntactic and pragmatic function assignment (in this ordering):

(13) $give_V\ (x_1\text{: John } (x_1))_{Agent/Top}\ (dx_2\text{: book } (x_2))_{Goal/Obj}\ (dx_3\text{: Bill } (x_3))_{Rec/Foc}$

After some *expression rules*, which (among other things) specify form and order of constituents as well as accent and intonation, the representation in (13) derives a sentence like (14) as an answer to a question like (15):

(14) *John gave the book to Bill*
(15) *To whom did John give the book?*

Concerning the data under the present discussion this would mean that the pronominal must somehow be assigned the pragmatic function *theme* - this would then account for the presupposition discussed in the previous section. It is assumed that the instantiation of a given phenomenon (i.e. pronominal insertion) serves to convey some sort of information which the absence of this phenomenon would not have.[13]

[13] Notice that the functions assumed in Functional Grammar are (partly) not treated as primitives within generative theory. Most generative linguists would agree that the notion of *subject* is a derived notion (formally defined as Spec/IP). The semantic function (i.e. theta-roles) are treated more controversially. If they are assigned at D-structure, then they would be considered as primitives. However there are attempts to derive the relevant interpretation from certain hierarchical configurations (similar to reducing subjecthood to a certain configuration rather than assuming it as a primitive). Finally it would be a desirable result if topic-focus would not be a

3.2. The Optionality Hypothesis

A formal (representational) theory claims that the relevant occurrence of a sentence is associated with a particular representation. Thus the representation differs whether the pronominal is there or not. Certain well-formedness conditions (on various levels of representation) are assumed for the relevant levels of representations. In our cases we find some well-formedness conditions on the level of domain D. Therefore we get the result that a given sentence is just wellformed in a certain context. It seems that the sentence under consideration serves to convey the relevant information. But this is just a byproduct of the representation assumed. As usual functional explanations are nothing more than assigning a function to a given phenomenon[14] which exists independently. In our cases the pronominal is associated with an index which has to be related to a previously established discourse referent. This is why the presupposition arises. The presupposition arises because the pronoun is anaphoric in nature and not because it is assigned a certain feature. It is therefore a byproduct of the 'intrinsic' anaphoric character of pronouns. We need not posit a feature (theme) to the pronoun and its correlated sentential element.

The next point to be considered is an implication the above line of reasoning leads us to. Assume that in a certain syntactic environment A the occurrence of a phenomenon X seems optional (again speaking about purely syntactic well-formedness conditions). As we have seen, this is just an apparent optionality once we take discourse into account. Assume furthermore that there is a syntactic environment B in which X is not optional but obligatory. According to what we have said at the beginning of this section, namely that any difference in form should imply a difference

primitive feature but rather follow from independent attested principles. This could be implemented by claiming that a particular position is defined as a focus position (i.e. FP) (as it is the case for subjects defined as SpecIP). Another possibility is given in Williams (1994) by claiming that presuppositions are the result of inherent anaphoricity.

[14] Cummings (1984) gives the following example for the problem with functional explanations in general. Suppose there is an ancient building that is ruined. A large stone fragment fell on a kind of zodiac mosaic and embedded itself there. No roof remains. Centuries later the ruin is visited by some tourists who think that this stone fragment was a sundial. Then the answer of a question like: '*Why is x there?*' (pointing to the gnomon of a sundial) might be '*Because it casts a shadow on the dial beneath, thereby indicating the time of day.*'. However, this answer 'explains' the presence of the gnomon by supplying a reason (of an agent) for putting it there. Therefore it might be assumed that this function is the reason of the gnomon's presence. Indeed the entity can be used to tell the time (which is also done by the people living close to the stone). Still the explanation for the existence of this entity should not be functional - the function is simply a byproduct, which can be assigned later.

in meaning it is expected that the occurrence of (a syntactically obligatory) X does not give rise to any presupposition.

Let us look at an example involving the phenomenon we are interested in here. Assume we find an instance where the occurrence of a pronominal in the matrix is not an option, but is syntactically forced. In this environment we expect that the occurrence of the pronominal does not give rise to the presuppositions discussed in section III.2. Consider the following example:

(16) (a) ?**weil mir gelungen ist den Fleck zu entfernen*
since REFL succeed is the spot to remove
'...since I was successful in removing the spot.'
(b) *weil es mir gelungen ist den Fleck zu entfernen*
since it REFL succeed is the spot to remove

Here we have a case where the occurrence of the pronominal is not an option. The sentence without *es* is considerably worth. I am not interested here in the exact reason for the obligatory occurrence of *es*. In the cases above it has to be noticed that the pronominal occurs in subject position. Its obligatory presence can presumably be reduced to the some version of the *Extended Projection Principle*. However, there are also instances where a pronominal in object position seems to be syntactically forced. This is for example true in the following sentence:

(17) *Ich hasse *(es) wenn mir das passiert*
'I hate *(it) when this happens to me'

The obligatory presence of the pronominal in this case might be reduced to the fact that a temporal clause cannot directly θ-identify an open position of the verb. However, the exact syntactic reason is not of any relevance here. What is however relevant is that the occurrence of *es* is syntactically necessary - i.e. the sentence without the pronominal is illformed independently of the (pragmatic) context.

According to what we have said in the above remarks we now expect that no presupposition needs to arise in this case. Thus we expect that the sentence with *es* is perfectly well-formed even in contexts where the occurrence of *es* results in ill-formedness in sentences where we find a syntactic option for the presence of *es*. This is a correct prediction:

(18) Q: *Was ist dir gestern gelungen?*
'What did you succeed in yesterday?'
A1: *Es ist mir gelungen, den Fleck zu entfernen.*
It is REFL succeeded, the spot to remove
A2: *Gestern ist es mir gelungen, den Fleck zu entfernen*
Yesterday is it REFL succeed, the spot to remove
A3: *?*Gestern ist mir gelungen den Fleck zu entfernen.*
Yesterday is REFL succeed the spot to remove

(19) Q: *Was haßt du?*
What do you hate?
A: *Ich hasse es, wenn mir das passiert*
'I hate it, if this happens to me'

The question answer pair above shows that the occurrence of *es* in the answer is possible (in fact it is again necessary since it is required by a syntactic well-formedness condition). I have given two possible answers since in case of (18 A1) we might be dealing with an instance of the so called *Vorfeld-es*. Thus their might be some intervening factors. I have controlled for this by using an answer like (18A2) where another constituent occupies SpecCP.

The following paradigms show the same phenomenon: if *es* is necessary for syntactic wellformedness conditions then the occurrence of *es* is not prohibited even in contexts where it is not allowed for 'optional' *es*.

(20) (a) *weil es mir gelungen ist, den FLECK zu entfernen*
since it REFL succeed is, the spot to remove
(b) *weil es mir sogar gelungen ist den Fleck zu entfernen.*
since it REFL even succeed is, the spot to remove

(21) (a) *weil ich es hasse, wenn es im FRÜHLING schneit*
since I it hate, when it in SPRING snows
(b) *weil ich es sogar hasse, wenn es im Frühling schneit*
since I it even hate, when it in spring snows

Focus can be associated with the extraposed constituent, indicating that we do not find the presupposition that the content of the sentential element (and of *es*) is already established. The generalization we can now draw is as follows:

(22) Optionality Generalization:

- *If an element X is syntactically optional then the presence of it induces some difference in meaning (in the above cases a 'pragmatic' difference). If it is present we get a reading A if it is absent we get a reading B.*
- *If the presence of the element X is syntactically forced then the meaning it has in case of its optional presence, does not necessarily arise: in this case it is ambiguous between the reading A and B.*

The problem we are faced with for an implementation of this generalization is as follows. It cannot be the case that the presence of the pronominal is associated with the same representation in every case. If it would be then the presupposition would also arise in the latter examples. But there it is just the case the presupposition (the necessary anaphoricity in the sense of Williams 1994) CAN arise but it need not. Interestingly this is exactly what Saussure says about differences in general: there can be a difference associated with a given item which does not reflect in the shape of the item itself but simply in the environment it occurs in:

> *"La valeur respective des pièces dépend de leur position sur l'échiquier, de même que dans la langue chaque terme a sa valeur par son opposition avec tous les autres termes."* (Saussure, 1915: 125f.)

Let me show that the generalization given in (22) is also found elsewhere, namely w.r.t. specificity. Several attempts to capture the notion of specificity result in the following generalization. Specificity seems to be connected to agreement. For instance it was observed that past participle agreement corresponds to a specific interpretation of the object. Again participle agreement in French is an instance of a phenomenon that was previously thought to be optional. It can however be shown that there is not really an option, but that the difference in syntactic realization corresponds to a difference in meaning. However, in case of agreement between the verb and the subject, where no such option is available (i.e. the verb has to agree with the subject) agreement does not necessarily correspond to a specific interpretation of the subject. This conclusion has also been drawn in Deprez (p.c.). Similar conclusions can be drawn from case-marking in Turkish as well as some agreement facts in Hindi.

I would like to briefly sketch a way of analyzing this generalization. According to the Principle of Full Interpretation every element needs some licensing. However if an element is syntactically optional then it is not necessary syntactically. If something

is not necessary then according to minimalist assumptions it shouldn't take place (or be present in case of insertion). Since the pronoun can occur although it is not syntactically licensed it must be licensed otherwise. Thus assume that in this case it is solely licensed at domain D. Thus I would like to suggest the following:

(23) The Optionality-Hypothesis[15]
- *If X is syntactically optional then the presence of X is represented (and thus licensed) at domain D.*
- *If X is syntactically obligatory then the presence of X can be expletive w.r.t. domain D.*

Concerning our examples this means that in case of a syntactically obligatory pronominal it can but need not have a representation at domain D. Thus the presupposition arises optional. One can get the presupposition but this is not necessarily so. On the other hand if the pronominal is syntactically optional (which means that it is syntactically expletive) it cannot be expletive w.r.t. domain D (since it needs to be licensed at some level of representation). In that case the presupposition necessarily arises.[16]

Notice that full lexical NPs acting as identifyees behave different. It seems to be the case that there is no syntactic option (i.e. the identifier is syntactically dependent on the nominal - either as a relative clause or an argument clause). However, although there is no syntactic option full lexical NPs always have to be represented at domain D in any case. Therefore full lexical NPs can never be expletive w.r.t. domain D.

[15] I have formulated the Optionality Hypothesis by explicitly mentioning domain D because this is the relevant generalization for the present discussion. However, more generally we could say that the presence of X must be licensed at least at one level of representation (thus instead of domain D it could also be LF, which I take to be a distinct from domain D.)

[16] This conclusion predicts that languages show morphological differences between the different pronominals, I.e. we expect that there are languages showing different pronominals corresponding to the two different uses in German. This is indeed the case as observed in Vikner (1993). He distinguishes between quasi-arguments and expletive subjects. To mention just a few instances: French has two different forms: *ce* and *il*, the former corresponds to a demonstrative and the latter to a non-demonstrative form. Similar facts are found in Italian, as well as some Germanic languages (like Norwegian, Swedish and Dutch). Many interesting issues would arise here especially concerning the difference between being expletive at domain D, or at LF. However, a detailed analysis would go beyond the scope of the present discussion.

3.3. Independent Evidence for the Optionality Hypothesis

Now we make a prediction concerning the data V&Z use as evidence for the possibility for a determiner to be expletive w.r.t. domain D. The prediction is as follows. Remember their claim that a definite determiner in French (and the same holds for German) can be expletive w.r.t. domain D. This is why the distributive inalienable interpretation can arise. In the cases they discuss the presence of the determiner was not a syntactic option (this is why it is ambiguous - it can but need not be expletive w.r.t. domain D). Now we predict that if we find an environment where the presence of the determiner is syntactically optional it cannot be expletive w.r.t. domain D. We can check this prediction with German *mass nouns* where the presence of the determiner is optional (as opposed to *count nouns*) as exemplified below:

(24) Mass nouns
- (a) *Peter spendet Geld*
  Peter donates money
- (b) *Peter spendet das Geld*
  Peter donates the money

(25) Count nouns
- (a) **Peter spendet Blume*
  *Peter donates flower
- (b) *Peter spendet die Blume*
  Peter donates the flower

We predict that the ambiguity in case of inalienable constructions discussed in V&Z arises only with count nouns but not with mass nouns. In the former case the determiner is optional. Therefore we predict that its presence must necessarily be represented at domain D. The inalienable (distributive) reading should disappear. This prediction is borne out:

(26) Count (inalienable) noun:
- (a) *Die Kinder haben die Hand gehoben*
  The children have raised the hand
- (b) **Die Kinder haben Hand gehoben*
  The children have hand raised

(27) Mass (inalienable) noun:
- (a) *Die Kinder haben das Blut gespendet*
  The children have the blood donated
- (b) *Die Kinder haben Blut gespendet*
  The children have blood donated

In case of count nouns (26a), where the determiner is syntactically obligatory there is an ambiguity between the distributive inalienable reading (*every child raised its hand*) and the alienable reading (*the children have raised one particular hand*). On the other hand in case of mass nouns (*blood*) which can have an inalienable interpretation there is no ambiguity (27a). The sentence can just mean that there was one particular amount of blood the children own and they donated it. The NP can not have the distributive inalienable interpretation. According to V&Z's analysis this means that the determiner cannot be expletive w.r.t. domain D. This is exactly the pattern which corresponds to the generalization in (22). An element X can be expletive at domain D (having no representation there). If it is however not licensed syntactically then it cannot be expletive at domain D.

If the Optionality-generalization turns out to be true then I think the Hypothesis in (23) is a seemingly plausible way to capture it - it provides indirect evidence for the analysis presented here. Standard assumptions would face the following problem concerning the implementation of the Optionality generalization: If a phenomenon X has the representation Y and Y is intended to capture the meaning effects then why do these meaning effects (associated with X) do not arise in case X is syntactically forced. This would correspond to saying that the actual representation of X is sensitive to the existence or non-existence to other constructions. Thus it would imply the possibility to compare different derivations, i.e. it would have to receive an optimality-like interpretation. On the other hand if one assumes some version of V&Z's analysis (i.e. the possibility for an element to count as expletive at domain D) plus the Principle of Full Interpretation one does not need such an optimality-like approach. This is so since if something is syntactically obligatory (i.e. licensed) then there must be some inherent property of a certain element which forces the presence of another element. It must then be possible to read off syntactic obligatoriness from one particular representation. On the other hand if syntax allows for an option, then there cannot be an inherent specification to force the presence of a certain element. This would then contradict some version of the Principle of Full Interpretation. Therefore the presence of this element necessarily needs to be licensed at some other level of representation (having to do with the interpretation). I will leave things here for further research but I think that the facts discussed so far can be taken as independent support for the analysis advocated here.

Now let us turn back to the heart of our analysis. If the representation of pronominals suggested here is correct it should also hold for other pronominals (not just the ones that are correlated with a sentential element). Combined with the assumption that discourse related phenomena are evaluated linearly in a step by step

manner (S-structure ordering being relevant) we expect some similar presuppositions w.r.t. other instances of pronouns. There are two empirical domains to look at: Right-Dislocation and backward pronominalization.

4. Right Dislocation

I will only briefly discuss instances of Right Dislocation. Consider the following example:

(28) *Maria hat ihm$_{X,*}$ schon gesehen, den Peter$_{*}$.*
Mary has him already seen, the Peter

Here we find a pronominal element in the matrix which is coreferent with a right-adjoined full lexical NPs that specifies its content. The relation between the pronominal and the lexical NP can again be interpreted as an instance of Identification. Again the pronominal is associated with an index X denoting the set of possible antecedents. Therefore the associated discourse referent must be already established, i.e. presupposed. The following question answer pair shows that this is a correct prediction:

(29) Q: *Wen hat Maria gesehen?*
who has Mary seen
A1: *#Maria hat ihn gesehen, den Peter.*
Mary has him seen, the Peter
A2: *Maria hat Peter gesehen.*
Mary has Peter seen.

The answer to a question like (29) cannot contain a pronominal. However if the discourse referent *Peter* is already available then (29A1) is a possible answer:

(30) Q: *Hat Maria den Peter schon gesehen?*
Has Mary the Peter already seen
A1: *Ja, ja, Maria hat ihn schon oft gesehen, den Peter*
Yes, yes, Mary has him already often seen, the Peter

The difference to other cases of Identification (involving sentential elements) is as follows. In order to pick the relevant discourse referent out of the set of possible antecedents the discourse referent itself (i.e. the proper name) is provided. In previous cases (or in case of relative clauses acting as identifiers) there the identifier provides

only a property which suffices to pick out the relevant referent.[17] I do not want to go into a detailed discussion of Right-Dislocation, since I do not have an analysis of right-adjoined NPs in general. I think that an adequate analysis for Right Dislocation would have to address the puzzle induced by the following paradigm.

(31) (a) *Ich habe (es$_i$) geglaubt, [$_{CP}$daß Hans gerne Bier trinkt]$_i$*
I have it believed that John likes beer drinking.
(b) *Ich habe ihn (da$_i$) schon gesehen, [$_{PP}$in dieser Bar an der Ecke]$_i$*
I have him (there) already seen, in this pub at the corner
(c) *Ich habe *(ihn$_i$) gesehen, [$_{NP}$den Peter]$_i$*
I have *(him) seen, the Peter.

On the one hand German allows for right-adjunction of NPs and CPs in case there is an overt correlate in the matrix. On the other hand if there is no overt correlate present in the matrix German only allows for right-adjunction of CPs and PPs but not of NPs. I do not have an account for this puzzle. The present section should just suffice to show that Identification is not only possible by means of a coreferent sentential element but also by a lexical NP.

5. Backward pronominalization

It is a well known fact that backward pronominalization induces some peculiar constraints on the interpretation of the pronoun. There are different views on how to interpret this fact. Here I will adopt the view that the denotation of the pronoun has to be presupposed (cf. Williams 1994).[18] Consider the following instance of backward pronominalization:

[17] The parallelism between these examples and the ones where a nominal is related to a sentential element suggests that the latter are also an instance of 'right-dislocation'. Cf. Wiltschko (1993) for a more detailed discussion concerning this conclusion.

[18] There are other suggestions found in the literature. For example V&Z claim that the relevant generalization is that the pronoun has to be referential. Notice that this is a strange claim especially given their assumption (discussed in section II.2.1.1.) that there is no level of representation where 'reference' plays a role. I will not go into a detailed discussion of the various approaches. The analysis presented here however supports the 'presupposition approach'. Moreover, I think that the notion of referentiality (as also noted in V&Z) is highly problematic. Thus even in their approach it should be translated as to saying that the pronoun has to denote. This however can be analyzed as saying that there must be an already established (i.e. presupposed) discourse referent. Remember that I assume (with Williams 1994) that 'presupposition' can be equated with anaphoricity.

(32) *Weil er zu viel Bier getrunken hat, wurde Hans ziemlich ausfällig.*
since he too much beer drunk has became John rather aggressive
'Since he has drunk too much beer, John got rather aggressive.'

It can be shown that the denotation of the pronoun has to be presupposed by means an already established test. Backward pronominalization is not legitimate if the following coreferent NP is focussed (focus indicating new information).

(33) **Weil er zu viel Bier getrunken hat wurde HANS ziemlich ausfällig.*[19]
since he too much beer drunk has became John rather aggressive

Notice that this phenomenon was first observed in Akmajian & Jackendoff (1970). They give the following English examples:

(34) (a) *After he woke up, John went to town*
(b) *After HE woke up, John went to town*
(c) *After he woke up, JOHN went to town*
(Akmakjian & Jackendoff, 1970: 124)

They note that coreference between *he* and *John* is just allowed in (34a) but not in (34b) or (34c) where either the pronoun or the name is stressed. This follows from the present analysis. On the one hand focus indicates new information. On the other hand I assume that backward pronominalization is just licensed in case the content of the pronoun is presupposed. Therefore it cannot be the case that the content of the pronoun is at the same time presupposed and conveys new information.

If it is true that backward pronominalization is only legitimate if the denotation of the pronoun is presupposed we have independent support for two aspects of the present analysis. First this phenomenon indicates that the pronoun cannot pick up its reference from the following coreferential NP. This fact is recently discussed in Williams (1994). He argues that it is really precedence rather than c-command which is responsible for dependency (extending this phenomenon to a range of other phenomena). The line of argumentation is as follows. If it would be the case that a pronoun cannot pick up its referent by a non-c-commanding NP then one would expect

[19] Notice that examples like (33) have been argued to indicate that focussed constituents move. The focussed NP would then be an operator, binding two variables (which can be ruled out by the bijection principle). Under such an approach these cases are assumed to be an instance of a weak cross over violation. However, it has been argued in various places that this cannot be correct. Among others cf. Williams (1994).

the same presupposition to obtain in case a name precedes the coreferential pronoun but still does not c-command it. This expectation is however not fulfilled as the following examples exemplify:

(35) (a) *Weil Hans zu viel Bier getrunken hat wurde er ziemlich ausfällig.*
since John too much beer drunk has became he rather aggressive
(b) *Weil HANS zu viel Bier getrunken hat, wurde er ziemlich ausfällig.*
since JOHN too much beer drunk has became he rather aggressive

If the coreferent name appears in a position preceding but still not c-commanding the pronominal element no presupposition arises, i.e. the content of the pronoun does not have to be presupposed. This is itself not a surprising fact since a pronoun can also be anaphorically related to an NP which does not occur in the same sentence. Inter-sentential anaphoric linking is possible:

(36) *Hans hat zu viel Bier getrunken. Deswegen wurde er ziemlich ausfällig.*
'John has drunk too much beer. That's why he became rather aggressive.'

However, if one assumes that the interpretation of a pronoun is evaluated after calculation of the entire sentence, then this fact cannot be explained. It should not make any difference whether the pronoun precedes or follows the name. Thus the facts concerning backward pronominalization are independent support for the claim that there is a step-by-step evaluation of (partial) DRSs. The linear ordering IS relevant for the interpretation, i.e. both, the Novelty-Familiarity Condition as well as the interpretation of pronouns is evaluated as soon as the relevant constituent is uttered.

Moreover, facts concerning backward pronominalization provide further evidence for the claim that no reconstruction effects show up for evaluating DRSs. The S-structure position (and not the D-structure position or the position we can reconstruct to at LF) is the relevant input for evaluating DRSs. Consider the following example:

(37) (a) *Welches Bild von ihm, hat Hans, geglaubt daß Peter gesehn hat?*
Which picture of him has John believed that Peter seen has
(b) *Welches Bild von Hans, hat er, geglaubt daß Peter gesehen hat?*
Which picture of John has he believed that Peter seen has

Here we are dealing with a wh-moved phrase containing either a pronoun (37a) or an R-expression (37b). As we have already seen wh-movement is subject to reconstruction. If the interpretation of (pro)nominals would be evaluated in the D-structure position we would expect these two sentences to have the same

presupposition. S-structure backward pronominalization should not play a role, i.e. we would expect that the denotation of the pronoun does not have to be presupposed. This is however not a correct prediction as the ill-formedness of coreference in case of stress on the R-expression indicates.

(38) **Welches Bild von ihm, hat HANS, geglaubt daß Peter gesehen hat?*[20]
Which picture of him has JOHN believed that Peter seen has

I conclude that we have independent support for the claim that there is a linear step by step evaluation of DRSs - a claim that is crucial for my analysis.

5.1. Backward Pronominalization in DRT

Given the above generalization that in case of backward pronominalization the content of the pronoun has to be presupposed it becomes again relevant to compare our approach to standard DRT assumptions. The following instance of backward pronominalization is discussed in Kamp & Reyle (1993) (henceforth K&R):

(39) (a) *A stockbroker who knows him likes Bill*
(b) *A stockbroker who knows Bill likes him* (K&R, 1993: 88)

The following DRS corresponds to the sentences in (a) and (b) respectively:

(40) (a)

x
stockbroker (x)
[x knows him]
[x likes Bill]

(b)

x
stockbroker (x)
[x knows Bill]
[x likes him]

(K&R,1993: 89)

[20] Notice that there is a reading available which allows for the coreferent reading even if the R-expression bears stress. Under this reading the denotation of the pronoun is presupposed but the R-expression is contrastively focussed. In this case an appropriate set of alternatives has to be present:
(i) *Welches Bild von ihm, hat HANS, *(und nicht Maria) geglaubt daß Peter gesehn hat*
Which picture of him has JOHN believed that Peter seen has

(41) *A stockbroker who knows him$_{x}$, likes Bill$_{x}$*[22]

(42) Pron: X = set of possible (male & human) discourse referents
x =?
$x \subseteq X$
Name: x = Bill

At the point of calculation where the discourse referent for *him* is evaluated all we know is that it must be contained in a set of possible discourse referents. (The descriptive content of the pronoun adds further restriction by means of the pronoun's (phi-) features: the relevant discourse referent must be male and human.) The presupposition arises that the discourse referent has already been established. However, at this point it might still be not clear which one of the available discourse referents is the intended one. Then, at a latter step, the index of Bill can be equated with the index of the pronoun (since Bill has enough descriptive content to pick out the actual referent).

This means that backward pronominalization and our cases of Identification (as well as Right-Dislocation) have a peculiar property in common: the (cataphoric) pronoun (re-)introduces an already established discourse referent. The representation of pronominals assumed here provides an account for the apparent contradiction: on the one hand pronouns cannot introduce a new discourse referent but on the other hand they seem to do so in certain environments.[23]

[22] Notice that I have not addressed the question concerning the appropriate indexing of proper names in chapter VII. I assume that they are also projecting an NP and a DP, the former being associated with X the latter with x. Assume furthermore that a name is basegenerated in N and moves to D (similar to pronouns). In this case X = x therefore no subset relation is induced. The discourse referent is directly picked out by the name. Since there is no set established names do not allow for restrictive modification. If a determiner occupies D (which is possible in German) the name cannot move there. In this case the name receives a property- (i.e.set) interpretation and restrictive modification (i.e. Identification) is rendered possible again as exemplified below:

(i) *Maria hat $[_{DP}$ Peter $[_{NP}$ $t_{[x]}]$, getroffen*
Mary has Peter met

(ii) *Maria hat $[_{DP}$ den $[_{NP}$ Peter$]_{[x]}$, den sie liebt getroffen*
Mary has the Peter who she loves met

[23] Notice that Williams (1994) solves this problem along the following lines:
"A definite pronoun in REST[rictive Clause] can start a new card: this card must be validated by the NUC[lear scope]; the validity of cards is checked after each File update, but does not hold of intermediate calculations." (Handout from GLOW 1994)
Since REST corresponds to (old) presupposed material he can explain the same effects. However, it is not clear what exactly REST and NUC correspond to and how this should be formally

Consider the description of the two DRSs given in K&R:

> *"At this point both DRSs (1.54)[=40a] and (1.55)[=40b] contain two reducible conditions. But now, to complete the two constructions so that the resulting DRSs represent the natural readings of (1.52)[=39a] and (1.53)[=39b] in which* ***him*** *is anaphoric to* ***Bill****, it is necessary to proceed differently in the two cases. In the case of (1.52) we must first reduce the condition representing the content of the main clause, so that we have a discourse referent for* ***Bill*** *when we reduce the (condition representing the content of the) relative clause, which requires that we deal with the pronoun* ***him****. In the case of (1.53) the order of reduction that produces the right DRS is the opposite: first reduce the relative clause and then the main clause."*
> (K&R, 1993: 89)

First the relative constructs its corresponding DRT in order to have a discourse referent when the main clause is translated into a DRS. Assuming this it would be a consequence that the pronoun can pick up its referent from a following NP. This is however contrary to what we have seen above - it misses the fact that the pronoun occuring in a position preceding the R-expression must be anaphoric to an already established discourse referent.[21] Standard DRT treatments for anaphoric resolution (which in case of backward pronominalization would be 'cataphoric' resolution) fail to capture a crucial property of backward pronominalization. However a theory that is intended to capture main properties of discourse should be able to capture these peculiar presuppositions.

5.2. Backward Pronominalization and Pronouns introducing discourse referents

Consider backward pronominalization under the light of the representation of nominals assumed here:

[21] As observed in Williams (1994) there is one peculiar exception to this generalization. The following pattern for anaphoricity within a sentence is assumed there:

*$[_{subord}$...pro...$]$ $[_{subord}$...antec...$]$
$[_{matrix}$...pro...$]$ $[_{matrix}$...antec...$]$
$[_{matrix}$...antec...$]$ $[_{subord}$...pro...$]$
$[_{subord}$...antec...$]$ $[_{matrix}$...pro...$]$

(Williams 1994: 7)

Dependence can be forward or 'backward' and down, but it must be subordinate. I have no explanation for this facts.

(45) (a) *Er$_i$ hat versprochen Maria das Buch zu geben, das Peter$_i$ gelesen hat*
He$_i$ has promised Mary a book to give which Peter$_i$ read has
(b) *Er$_i$ hat Maria das Buch zu geben versprochen, das Peter$_i$ gelesen hat.*
He$_i$ has Mary a book to give promised which Peter$_i$ read has

At least for German these data suggest that we are dealing with a subject-object asymmetry depending on the pronoun that induces the condition C effect.

Notice that similar examples have traditionally been used to determine the adjunction site of extraposed constituents. It has been argued that the following paradigm is indicative for the relevant adjunction site:

(46) (a) *I sent her$_i$ many gifts last year that Mary$_i$ did not like*
(b) **I sent her$_i$ many gifts that Mary$_i$ did not like last year*
(c) **She$_i$ was sent many gifts last year that Mary$_i$ did not like* (C&R, 1990: 29)
(d) **It bothered her$_i$ that Rosa$_i$ had failed* (Reinhart 1983: 49)

Condition C effects are argued to be induced just in case of extraposed argument CPs but not in case of extraposed relatives. On basis of these examples it is argued that extraposed relatives are adjoined higher than argument CPs. If true then my analysis is faced with a problem. Notice that my analysis (i.e. the locality constraint I am assuming) predicts exactly the opposite: the subject CP should occur higher than the relative clause since its antecedent occurs higher in the matrix than the head of the relative clause. As pointed out to me by Edwin Williams the judgements in (46) are problematic in itself. Consider the contrast in (46a-b). My analysis suggests that there is no such contrast (and it seems that this contrast is indeed questionable). Traditional analyses are however confronted with an empirical problem (first noticed in Haider (1993)[26], having to do with the restriction on multiple extraposed elements. Notice that I assume that the adjunction site of the extraposed constituent is determined by the position of the correlate in the matrix: the identifier has to be adjoined to the first maximal projection dominating the identifyee, i.e. hierarchically the identifier has to be higher than the identifyee. If two constituents are extraposed we find exactly the ordering restriction predicted by my analysis.

[26] Haider claims that there is a restriction on multiple extraposed constituents to the effect that the relative clause has to precede the argument clause. However, in his examples this is true only because the correlates in the matrix occur in positions that necessarily result in this particular kind of ordering.

5.3. Backward Pronominalization and Condition C effects

As discussed above there is a peculiar licensing requirement for backward-pronominalization. The (cataphoric) pronominal is only licensed if it re-introduces an already established discourse referent. This is however not allowed in any arbitrary position. It seems to be the case that on the one hand the nature of the clause the pronominal occurs in plays a role[24] and on the other hand it makes a difference whether the pronominal itself is a subject or an object. For the latter distinction consider the following example from Akmajian & Jackendoff (1970: 125):

(43) (a) *What bit him$_i$ was the dog Mary had warned John$_i$ not to buy.*
(b) **What he$_i$ kicked was the dog John$_i$ told Mary not to go near.*

Without going into any detail I would just like to mention what this could mean for some other well-known phenomena relevant for extraposition. It has been observed in Wiltschko (1994) that there is a peculiar subject-object asymmetry concerning apparent condition C effects. As exemplified below an pronominal object does not induce condition C effects no matter whether the coreferent R-expression is in the c-command domain of the pronoun or not:

(44) (a) *Peter hat sie$_i$ gebeten ein Buch zu lesen, das Maria$_i$ noch nicht kannte*
Peter has her$_i$ asked a book to read which Mary$_i$ yet not knew
(b) *Peter hat sie$_i$ ein Buch zu lesen gebeten, das Maria$_i$ noch nicht kannte*
Peter has her$_i$ a book to read asked which Mary$_i$ yet not knew
(c) *Peter hat sie$_i$ gebeten ein Buch, das Maria$_i$ noch nicht kannte zu lesen*
Peter has her$_i$ asked a book which Mary$_i$ yet not knew to read
(d) *Peter hat sie$_i$ ein Buch das Maria$_i$ noch nicht kannte zu lesen gebeten*
Peter has her$_i$ a book which Mary$_i$ yet not knew to read asked

The most relevant example is the one in (44d), where the argument CP and the non-extraposed relative are clearly in the c-command domain of the pronoun and still coreference is possible.[25] However, subject pronouns do induce ungrammaticality in examples similar to the ones above:

implemented.

[24] This is for example observe in Williams (1994) as mentioned in Fn. 21.

[25] A similar phenomenon is found in English:
(i) *?Peter asked her$_i$ to read a book that Mary$_i$ didn't know*
(ii) **He$_i$ promised Mary to give her the book that Peter$_i$ wrote* (Grimshaw, p.c.)

(47) (a) *It$_i$ struck a grammarian$_j$ last month [who$_j$ analyzed it$_i$][that this clause is grammatical].*

(b) **It$_i$ struck a grammarian$_j$ last month [that this clause is grammatical]$_i$ [who$_j$ analyzed it$_i$]*

(48) (a) *Es$_i$ ist einem Grammatiker$_j$ aufgefallen [der$_j$ das untersuchte]$_i$ [daß dieser Satz ungrammatisch ist]$_i$.*

(b) **Es$_i$ ist einem Grammatiker$_j$ aufgefallen [daß dieser Satz ungrammatisch ist]$_i$ [der$_j$ das untersuchte]$_i$.* (Haider 1993)

The examples above are intended to show that the argument clause has to be higher then the relative clause, so the condition C effect cannot really be taken as a test for the height of the adjunction site. Notice that we do find Condition C effects even with multiple extraposed elements:

(49) (a) *Es$_i$ hat ihm$_k$ jemand$_j$ prophezeit [von dem$_j$ sich Max$_i$ wahrsagen ließ]$_k$ [daß Max$_i$ ein Schloß erben würde]$_k$*

(b) *Someone$_j$ has prophesied her$_k$ [who Mary$_i$ met at a party]$_k$ [that Mary$_i$ will inherit a castle]$_k$* (Haider 1993)

So we seem to be faced with a problem which apparently cannot be reduced to a structural account. If the generalization given above is on the right track we might draw the following conclusion which could be a way to solve the problem mentioned in Haider (1993). Let us adopt the view that there is no Condition C of binding theory.[27] Notice that condition C is not formulated in the way licensing conditions in general are defined. It is rather defined in the opposite way, namely that an element must not be bound, i.e. licensed. This is however not in the spirit of the general theory of licensing as discussed in chapter 1. Notice moreover that all the examples above involve a pronoun linearly preceding a name, thus in a way we are dealing with instances of backward pronominalization. We might assume that the examples traditionally argued to be instances of condition C violations are really instances of illegitimate backward pronominalization. I will not pursue any further implication of such a view but rather leave this line of reasoning as a matter of future research. What is however crucial for the present analysis is the conclusion that condition C effects cannot be used as a reliable test for the adjunction site of extraposed constituents.[28]

[27] See for example Reinhart (1983) arguing that apparent Condition C-effects are to be analyzed as violations of independently motivated pragmatic constraints.

[28] Notice that data involving extraposition (or Heavy NP Shift) in general show a pattern that leads to the following conclusion: as for constituency they behave as if they were right-adjoined whereas for some other phenomena (like condition C-effects or polarity binding) they

6. Accessibility for anaphoric resolution

Any theory dealing with anaphoric dependencies needs to be restrictive enough to explain what can count as a possible antecedent. Obviously such a theory is also crucial for the analysis suggested here. I have argued that a pronominal element (both, in case of Identification as well as for backwards pronominalization, which the former is a particular instance of) can (re-)introduce a discourse referent which has already been introduced. However, for some reason this discourse referent seems to be not immediately accessible for the pronoun. Thus we need a theory of when and why a discourse referent is or is not accessible for anaphoric linking in general - we need a general theory of accessibility. In this thesis I cannot present a fullfledged theory of accessibility - especially concerning the accessibility within discourse such a theory would need a very fine-grained analysis concerning the (hierarchical) structuring of discourse representation. Lots of work still needs to be done. Such a theory is partially developed is Asher (1993). I will therefore rely on his assumptions. Moreover, I will only look into some selected phenomena, others have to be neglected, although I think that many interesting issues would arise.

6.1. Accessibility within the Theory of Grammar

Let me start with some remarks on the theory of accessibility in general. In doing so I will make a side-step to a different domain, where the notion of accessibility is relevant. This is the domain of morphology. Consider the following examples:

(50) (a) **Klara hat einen guten [Kraut]salat gemacht. Sie hat es$_i$ frisch vom Garten geholt.*
'Claire has prepared a good cabbage salad. She has just picked it up from the garden.'

(b) *Klara hat aus [Kraut]$_i$ einen guten Salat gemacht. Sie hat es$_i$ frisch vom Garten geholt.*
'Claire has prepared a good salad from cabbage. She has just picked it from the garden.'

behave differently (as if they were more deeply embedded, arguing for a right-branching analysis). This phenomenon has been analyzed in various ways (for example Pesetsky (1990) argues for two different structures assigned to the relevant data. Broxly (1994) argues that there is a genuine difference between dependency and non-dependency relations in addition to position reconstruction effects).

In general there are several ways to explain the fact that an element (or a feature) appears to be higher as it really is. The most common assumptions are percolation or inheritance for morphology (cf. among others di Sciullo and Williams 1987); Lieber (1983)) and (LF-)movement for syntax (cf. among others May 1985). Notice however that Riemsdijk & William's (1981) account for LF-phenomena is in terms of feature percolation rather than movement. On the other hand Pesetsky (1985) claims that there is LF-movement within morphology. What these approaches have in common is that both a percolation approach and a movement approach rely on hierarchical constituency, i.e. certain hierarchical configurations allow for movement or percolation and others do not. If we turn things around we can say that the impossibility for anaphoric linking is indicative of some notion of inaccessibility. Within the syntactic and the morphological domain inaccessibility is reduced to restrictions that are governed by hierarchical structure. Thus we might assume that inaccessibility for anaphoric linking within discourse is an indication of hierarchical structuring of discourse. If true we can conclude that hierarchical structure (resulting in constituenthood) is needed not only in syntax and morphology but also in the domain of discourse. As we will see there are constraints on the accessibility for anaphoric resolution within discourse. First let me briefly discuss anaphoricity within syntax, and by that we will compare binding and Identification at domain D.

6.2. Anaphoricity in syntactic configurations: Binding Theory

Let us look at an area where 'anaphoricity' plays a role within syntax: binding theory which is assumed to consist of three principles:

assumed that X associated with N is not available any longer. What is particularly interesting is the following observation. On the one hand it is assumed that there are various different relations possible between the head and the non-head of a compound. However, one particular relation is not an option, namely the concessive. Thus a compound like:

(iii) *Feueranzug*
firesuit

can not denote a suit, that is a suit despite fire. On the other hand Identification (in our sense) is in principle possible w.r.t. various relations (as we will see in chapter VIII.) with one particular exception namely the concessive.

(iv) *Maria hat (*trotzdem) Bier getrunken, obwohl Hans es ihr verboten hat.*
Mary has (despite) beer drunk, although John it her forbidden has

Concessive clauses do not allow for an overt correlate in the matrix. This could suggest that concessive Identification is neither possible morphologically nor at domain D.

The two pieces of discourse differ in the following way. In (50a) we find the compound *Krautsalat* ('cabbage salad'). In this case it is not possible to anaphorically refer to the noun *Kraut* ('cabbage'). On the other hand this noun occurs as an independent lexical item (embedded in a PP) in (50b). Here it is accessible for anaphoric linking. The pronoun *es* ('it') can refer to *Kraut* in (50b) but not in (50a). An element below the X°-level is not accessible for anaphoric resolution. This is a well known fact. The subparts of a complex word are not visible within syntax. However certain phenomena of word-formation seem to violate this generalization. One example concerns information about *argument-structure*, which can still be available even if it is a feature of a subpart of the complex word (which is not even the head of the noun):

(51) *a baker of bread*

In the example above the thematic structure of the verb is still available for the derived nominal. Another instance of this phenomenon involves syntactic modification which seems to apply only to a subpart of the complex word:

(52) (a) *ein starker Raucher*
'a strong smoker'
(b) *Es handelt sich wahrscheinlich um eine Männerleiche*
'We are presumably dealing with a male-body'

In the above examples we find an adjective (52a) and an adverbial (52b) respectively. Modification here is ambiguous as to whether they modify the whole complex word (*Raucher*, *Männerleiche*) or just a subpart of it (*rauchen*, *Männer*). This phenomenon induces the need for finding well defined restrictions for when and how far various features can percolate. It is necessary to determine which features of the compound are visible in syntax.[29]

[29] Notice that some properties of compounds are also relevant for the present discussion concerning Identification. My analysis crucially derives the obligatory rightness of extraposition from a precedence constraint (derived by the Novelty Condition). However consider the following example:

(i) *Peter hat das [daß Maria kein Bier trinkt]-Argument präsentiert.*
Peter has the [that Mary no beer drinks]- argument presented

If the potential identifier precedes the noun it is interpreted like a 'phrasal-compound', similar to the following example:

(ii) *die 'rühr-mich-nicht-an-Einstellung' Marias*
the 'leave-me-alone-attitude' Mary's

The non-head of the compound can be interpreted as an identifier but in a special way. It has to be noticed that the potential identifier still does not precede x (associated with D°), the index for the actual discourse referent. Since compounds in general induce some sort of opacity it can be

(53) Binding Theory
A: An anaphoric element has to be bound within a local domain X
B: A pronominal element has to be free within a local domain X
C: An R-expression has to be free

It seems that we are dealing with some sort of *hierarchy* concerning the descriptive content of a nominal element.[30] Anaphors are at the lowest rank. It is their lexical property that they cannot denote, i.e. they need a syntactic antecedent which helps to identify their content. Similar effects can be observed w.r.t. Identification at domain D. In section II.2.4. we have seen elements which are lexically specified for the need of Identification (the complex determiner *derjenige*). Pronouns are ranked higher on this hierarchy. They differ from anaphors in that they can denote, however their inherent denotational content does not suffice in order for them to pick up their referent without some sort of antecedent. The only restriction on their denotation they impose is some phi features (like gender and number). This lack of ability to inherently denote is encoded by claiming that they need some sort of discourse antecedent. They cannot introduce a new discourse referent.[31] However, they can also be identified but Identification is not necessary for syntactic wellformedness. The need for Identification is dependent on the availability of the discourse referent.

Lexical NPs (and proper names) are on the highest rank of the hierarchy concerning the ability to denote. Here we find restrictions within NP (by means of the descriptive content of N as well as modifying elements contained within NP). However, as we have seen lexical NPs can also lack enough descriptive content to actually pick out the relevant discourse referent. This is the most crucial point of the analysis presented here, i.e. lexical NPs as well as pronominals can be identified at domain D.

[30] Empty elements are obviously on the lowest end of this hierarchy, i.e. they do not have descriptive content at all and in fact all of them need some sort of Identification. Remember that many analyses for licensing of empty elements distinguish between formal licensing and licensing w.r.t. their content, i.e. Identification.
Notice moreover that Lasnik (1991) defines binding theory in terms of such a hierarchy:
(i) *A less referential expression may not bind a more referential one.* (Lasnik, 1991: 19)

[31] The only exception are deictic pronouns. They can introduce a discourse referent by virtue of the fact that the discourse referent is (physically) present.

6.3. Syntactic antecedence vs. discourse antecedence

Binding theory and anaphoric relations in general are two different relations. As binding theory indicates the accessibility of a given element X for 'binding' another element Y is heavily determined by hierarchical structure. Binding is defined as follows:

(54) A binds B iff A and B are coindexed and A c-commands B

Moreover there is a given domain (also defined hierarchically) which counts as the binding domain. Anaphors must be bound within this domain, pronouns must not be bound within this domain but may be bound by an antecedent outside their binding domain. R-expressions may not be bound at all. This sort of anaphoric relation (referred to as binding) differs from anaphoric relation within discourse (i.e. intersentential anaphoricity) - i.e. c-command cannot be at play for discourse antecedence.

A similar point is made in Williams (1994). He claims that the pattern of anaphora is a pattern of allowed *dependence* (which is governed by precedence). The pattern of permitted *coreference* on the other hand is to be equated with binding theory (governed by c-command). As evidence for the claim that the two notions are distinct he gives the following examples. On the one hand it can be shown that dependence can be barred where coreference would be permitted:

(55) **Anyone$_i$ can try I_S to hand it$_i$ in to me$_j$ who has written HIS$_i$ TERM PAPER$_i$*

We are already familiar with these examples from the discussion of backward pronominalization: although the pronoun does not c-command the R-expression coreference is still blocked. On the other hand there are examples where coreference is barred where dependence is permitted (originally discussed in Williams (1993):

(56) *Every hospital administrator hates the hospital*

Here *hospital* is analyzed as being dependent on *every hospital administrator*.

> *"in the sense that its referents vary according to the values assigned to the variable bound by the universal quantifier."* (Williams, 1994: 8)

Coreference between the subject and the object is blocked whereas dependence is allowed.

For the present discussion I would like to combine things in a slightly different way. Discourse anaphoricity in our terms is related to linear ordering (precedence) (in addition to the hierarchical structuring of discourse representation) and binding is related to hierarchical (sentence) structure. One effect of this difference was already mentioned namely that binding is subject to reconstruction effects whereas discourse reference is not. This has to do with the fact that (syntactic) binding is dependent on syntactic structure and moreover it is calculated at LF (i.e. not dependent on the linear ordering or S-structure position). Discourse dependency however is dependent on S-structure linear ordering. Within this domain reconstruction is not relevant.

Let us refer to binding (Williams' coreference) as *syntactic antecedence* and to discourse anaphoricity (Williams' dependency) as *discourse antecedence*. As we have already seen from the discussion of Williams (1994) both phenomena are related though they do not map in a one to one relation nor is the former a subclass of the latter. Thus we find the following situation with all 4 possible combinations:

(57) A given constituent can be
1) syntactically available and discourse unavailable
2) syntactically unavailable and discourse available
3) syntactically available and discourse available
4) syntactically unavailable and discourse unavailable

The first case is exemplified by William's example (55) (coreference possible and dependency blocked). Another instance of this (in our terms being syntactically available but not discourse available) would be a bound variable reading. Thus consider:

(58) *Es gibt keinen Frosch, der sein Revier nicht kennt. #Er geht dort jeden Tag.*
'There is no frog, who does not know his district. #He walks there every day.'

Although the pronoun embedded in the relative clause is anaphorically related to the negative quantified NP *kein Frosch* ('no frog') the pronoun in the second sentence cannot take *kein Frosch* as its antecedent. Since intersentential anaphoricity must be discourse anaphoricity we can conclude that *kein Frosch* in the first sentence is not accessible for discourse anaphoricity. Notice however, that being not accessible for discourse anaphoricity does not necessarily imply that no discourse referent is induced.

(This will become important when we look at instances of Identification of quantified and predicative NPs).

The second case (being syntactically unavailable but discourse available) is instantiated in case an antecedent A does not c-command (i.e. does not bind) another element B. Thus we find:

(59) *Weil Peter zu viel getrunken hat, wurde er ziemlich ausfällig. (Er hat Maria angemacht.)*
Because Peter has drunk too much, he became rather aggressive. (He abused Mary.)'

The pronoun in the matrix clause is not bound by the R-expression in the topicalized adverbial clause but it can still be coreferential. Syntactic unavailibility but discourse availability is also instantiated by certain kinds of binding violations. Binding theory can rule out a given representation whereas discourse antecedence could still be possible. Consider an example of a condition B violation:

(60) **Hans liebt ihn. (Er ist ein narzistisch veranlagter Mensch.)*
*John loves him. (He is a narcissistic disposed man).

Binding theory rules out this sentence although discourse antecedence should be possible. This is evident for various reasons. First, an appropriate anaphor is possible (i.e. the reflexive). Secondly, the sentence is comprehensible even under the coreferential reading. Thirdly the pronominal in the subsequent sentence is perfectly wellformed under the coreferential reading. The third case (syntactic and discourse availability) is instantiated in most cases of binding a pronoun (an exception being the bound variable reading of pronominals as discussed above):

(61) *Maria hat gesagt, daß sie Hans verlassen wird. (Sie hat auch schon alles in die Wege geleitet).*
Mary said that she will leave John. (She already prepared everything)

The antecedent NP *Maria* is in a position where it can c-command the pronoun. Therefore it can be argued that the pronominal is syntactically bound by the R-expression. Consider now the fourth case, i.e. being neither syntactically nor discourse accessible. Here we come into an area which especially interests us here. We have to distinguish between the nature of the possible antecedent on the one hand and the nature of the anaphoric element on the other hand. Abstracting away from bound variables we can say that any antecedent which is not discourse accessible is also not

Definition:

Let K be a DRS, x a discourse referent and γ A DRS-conditions. We say that x is *accessible from γ in K* iff there are $K_1 \leq K$ and $K_2 \leq K_1$ such that x belongs to U_{K1} and γ belongs to Con_{K2}.[33] (K&R, 1993: 120)

This definition for discourse accessibility is reminiscent of the constraints that hold for binding in syntax: the antecedent must be within a superordinate DRS (which can be compared with c-command) and within a local domain.

There is also a constraint concerning the nature of the possible antecedent.[34] It seems as if proper names are always accessible (given that they are salient enough). They always introduce discourse referents which 'climb up' as high as accessibility needs them to. This is not a surprising result: As is well known it is also true for purely syntactic phenomena like scope assignment (proper names always behave as if they have widest scope). Certain kinds of other elements behave exactly in the same way. This is the case for some predicative DRSs. They characterize abstract entities which gives rise to a term interpretation. Thus those DRSs function like names[35]. However, this is true just in case the discourse referents within the predicative NPs are fully accessible themselves. Asher (1993) gives the following examples to exemplify this fact:

(62) (a) *Everyone believes that Fred should not allow [unsafe flying in his airplane]$_i$. It$_i$ led to an accident before and it$_i$ will lead to a accident again.*

(b) *John didn't see [Mary leave]$_i$ and Sam didn't see it$_i$ either.*

[33] Here ≤ refers to *weak subordination* (the relation which holds between K_1 and K_2 if either K_1 is identical to K_2 or K_1 is subordinate to K_2. Subordination is defined as follows:
Definition:

(i) K_1 is *immediately subordinate* to K_2 iff Con_{K2} contains the condition $\neg K_1$.

(ii) K_1 is *subordinate* to K_2 iff either

(a) K_1 is immediately subordinate to K_2, or

(b) there is a DRS K_3 such that K_3 is subordinate to K_2 and K_1 is immediately subordinate to K3 (K&R, 1993: 119)

[34] cf. chapter VII. for a discussion of another constraint for the nature of the antecedent, which is referred to as *Novelty Condition* in Wasow (1972) (not to be mixed up with Heim's Novelty Condition). It basically says that an antecedent must not have less descriptive content as the anaphoric element.

[35] This property of predicative NPs will be supported by some facts concerning Left Dislocation. Thus we will see in chapter VII. that names and predicative NPs behave similar w.r.t. the possibility to be left dislocated. This distinguishes them from indefinite NPs.

syntactically accessible. Thus we can simply look at intersentential anaphoricity. In principle there are two different ways (which both seem necessary) to look at accessibility. Both depend on a hierarchical organization of discourse.

On the one hand it seems to be necessary to determine a *salience* ranking of discourse referents relative to some other discourse referent z that receives its interpretation by means of anaphoric resolution. In most cases this corresponds to a ranking of possible discourse antecedents. However, this is not always the case. Salience is a matter of degree: the greater the degree of salience of a discourse referent x relative to an anaphorically related discourse referent z, the more likely is the process of anaphora resolution to identify x with z. Moreover (as we will see in subsequent sections) this degree interacts with the hierarchy concerning the ability to denote. The more salient a given discourse referent is the less (inherent) denoting properties need the anaphorically related element have and vice versa: the less salient a discourse referent is the more denoting properties the anaphorically related element must have.

The other necessity for hierarchical organization of discourse is implied by the fact that not all discourse referents ARE accessible to anaphoric resolution. Here principles have to be found which restrict anaphora resolution in general. This second sort of hierarchical organization is akin to the nature of accessibility within syntax, it is configurational. Consider the following quote from K&R and their definition of Accessibility which is defined hierarchically[32]:

> *"...whether the relation holds between, say, a pronoun π and a discourse referent x depends on the one hand on the DRS which contains the syntactic tree in which π is a constituent as one of its conditions, and on the other hand on the DRS whose universe contains x. Whether x is accessible to π is fully determined by the configurational relationship between these two DRSs.* (K&R, 1993: 118)

[32] For completeness let me give Asher's (1993) definition of discourse accessibility:
Definition (**Accessibility**): Let $K = K_n$ or K be subordinate to K_n. Then:

(i) If $x, y \in U_K$, then y is accessible to x in K_n

(ii) If K' is subordinate to K and $x \in U_{K'}$ and $y \in U_{K'}$, then y is accessible to x in K_n

(iii) if $y \in U_{K'}$, $x \in U_{K''}$ and $K' \Rightarrow K''$ is a complex condition of K, then y is accessible to x in K_n.

(iv) otherwise, y is not accessible to x in K_n.

"Accessibility is a relation on discourse referents in the universe of DRSs. Roughly, the constraint says to find all the accessible discourse referents accessible to some discourse referent x, one must look in those universes of DRSs to the left or superordinate to the universe in which x is declared."
(Asher (1933: 76f.)

(c) *Mary saw [no one leave the building]$_i$ and Sam saw it$_i$ too.*
(d) **Mary saw [Fred hit no boy]$_i$. Jean saw it$_i$ too.*
(e) *??Mary saw [Fred give no present to a child in the orphanage]$_i$. Sam saw it$_i$ too.*

Asher's (1993) explanation is as follows: Although the event type occurs within the scope of a negation operator, a quantifier and an attitude operator in (62a) and inside a negation operator in (62b) the anaphoric connections intended in (62a-b) are perfectly felicitous. The event types with their definite arguments are accessible to any subsequent discourse referent. On the other hand the oddness of the examples in (62c-e) is explained by the assumption that the accessibility of SubDRSs (i.e. the predicative DRS which count as entities in the model) is limited by the accessibility of the discourse referents that occur in them.

As can be seen on basis of this few remarks, an analysis of discourse structure and subsequently the definition of discourse accessibility is rather complex. In the next section I will show that our cases of Identification (which are related to anaphoric relations relate to some general notion of (discourse) accessibility.

6.4. Anaphoric Resolution within DRT

For completeness let me discuss the DRT-implementation of intersentential anaphora since we find a parallelism to the present analysis of Identification. Consider the following discourse:

(63) *A boy kicked Fred. He cried.*

The translation of the second sentence, which involves a pronoun that is anaphorically related to a nominal NP in the first sentence introduces the following partial DRS with subsequent construction of the DRS of the second sentence:

(64) λP

z
P(z)
z = ?

z
cry (z)
z = ?

The content of the variable z is not established. It is however assumed that there has to be a discourse referent (in U_K) to which this pronoun corresponds. Within DRT a condition like $z = ?$ is an *incomplete condition*. Anaphoric resolution comes into play now. It completes the incomplete condition. The DRS in (64) has to be updated with the DRS of the first sentence. This is done by means of taking the union of the universes and the condition sets of the two DRSs (where $\underline{\cup}$ stands for the operation of DR-theoretic union) resulting in the DRS given in (66):

(65) DEFINITION: DRS-update $(K_1, K_2) = \langle (U_{K1} \cup U_{K2}), (Con_{K1} \cup Con_{K2}) \rangle = K_1 \underline{\cup} K_2$

(66)

x,y,z
boy (x)
Fred (y)
kick (x,y)
cry (z)
z = ?

(Asher 1993: 73)

Anaphora resolution will then find an appropriate discourse referent (other than z) which replaces ? in $z = ?$. In the example above it will identify y with z. Accordingly the following completed DRS is created:

(67)

x,y,z
boy (x)
Fred (y)
kick (x,y)
cry (z)
z = y

Notice that this sort of phenomenon is similar to Identification. There is an intended discourse referent, the content of which is not clear at one step of the derivation (symbolized as $x = ?$). After updating (by means of taking the union of the two DRS) the content of x is identified with a (previously established) discourse referent. This means that anaphora resolution and Identification work in a similar way.

However, not every discourse referent can be identified with an incomplete condition. The intended discourse referent must be salient enough. However, the notion of salience does not suffice as we will see. I will pick out one particular instance where a given (possible) antecedent is not accessible but still salient, (i.e. in the field of attention). This involves the operation of *summation*, as for example in case of

coordination. The reason that the pronoun is not able to pick up an (already established) antecedent has to do with the fact that once summation has taken place (i.e. once a set is created) the individuals within this set are not available anymore. (Notice that I do not imply that this is the only phenomenon of inaccessibility for anaphoric linking.) We will furthermore see that properties of the anaphoric element (for example whether it is a pronominal or a definite description) heavily influences accessibility for anaphoric resolution. We find an interaction between the descriptive content of the anaphoric element, salience and accessibility respectively.

6.5. Accessibility, Prominence and Salience

Accessibility (as defined above) is not the only condition which is important for anaphoric relations. Recall that according to what we have discussed so far both (definite) pronominals as well as definite DPs are inherently anaphoric, i.e. both need a (previously established) discourse antecedent (given the Familiarity Condition). However there is an inherent difference between definite DPs and pronominals - as noted in Heim (1982). Definite DPs CAN introduce a new discourse referent by means of *accommodation* (a mechanism originally proposed in Hawkins 1978). As an example consider the following case of *associative anaphoric use* of a definite DP:

(68) *John read [a book about Schubert]$_i$ and wrote to the author$_i$*
(Heim, 1982: 371)

The possibility to use a definite DP in this case (apparently violating the Familiarity Condition) is argued to be due to the possibility to accommodate the new discourse referent: it has to be linked by cross-reference to some already present discourse referent (in (68) *the author* is linked to *a book about Schubert*). Furthermore Heim (1982) argues that pronominals cannot be saved by accommodation (with some exceptions[36]) because they are subject to the following condition:

(69) Prominence Condition: For a pronominal definite NP$_i$ to be felicitous w.r.t. a file, i must be a prominent element of Dom(F). (Heim,1982: 385f.)

[36] An exception Heim discusses involves so called *paycheck* sentences (cf. Karttunen 1969) and so called *pronouns of laziness*:
(i) *Every wise man gives his paycheck to his wife. John gives it to his mistress.*
(for discussion cf. Heim, 1982: 387f.)

"What does "prominent" mean? Let us assume that a file is not just an amorphous bunch of cards, but is organized in such way that a small number of cards enjoy a privileged place, "on top of the file", so to speak. These are always the cards that the file clerk had to handle most recently, i.e., that were most recently introduced or updated."
(Heim, 1982: 386)

On basis of the following examples it is furthermore argued that prominence cannot be saved by accommodation:

(70) (a) *John has a cat and a dog. The cat's/#it's name is Felix.*
(b) *John is married. His wife/#she is nice* (Heim, 1982: 384)

Another instance of definite DPs being familiar without explicitly introducing the discourse referent (by means of an indefinite DP) is referred to as (*contextual* or *perceptual*) *salience*. In that case relevant referent is in the background knowledge of the participants in the discourse. Those (potential) discourse referents are however not immediately accessible as antecedents - in order for them to be available they must be rendered salient. The notions of salience, accessibility and prominence seem to interact both for anaphoric as well as cataphoric relations. The felicity of pronominals or definite descriptions in the examples in (70) seems to be a matter of descriptive content. According to the representation of (pro)nominals I am assuming the above pattern follows: in any case X (the index of NP) refers to the set of possible referents. These are however defined relative to the context of discourse. They should include contextual salient potential discourse referents as well as explicitly introduced discourse referents (which might be not prominent at the time of utterance). The nominal is felicitous only in case it provides enough restrictions for the relevant discourse referent to be picked out. Lexical DPs always come with some restriction (provided by N) whereas pronominals only have phi-features as a restriction. We therefore expect the following situation: if a definite DP is used that dominates an N denoting a property that is satisfied by both individuals contained in the set of previously established referents, then anaphoric resolution should again be blocked. This is a correct prediction as can be observed below:

(71) *John has [[a cat]$_i$ and [a dog]$_j$]$_k$. #[[The animal's]$_{i,j}$ name is Felix.*

The definite DP in the second sentence contains an N (*animal*) that denotes a property which both of the possible antecedents satisfy. Therefore it cannot be successfully linked to either of the two discourse referents. This means that there is no clearcut two-way distinction between pronouns and definite DPs. Such a distinction is however

implied by Heim's Prominence condition as well as by the following quotation of Vallduví:

> *"Actually to be exact, there is a further difference between definites in general and pronouns in particular. Both denote preexistent addresses but differ in that pronouns denote salient preexistent addresses and other definites nonsalient ones. In other words, definites trigger an activation of a dormant preexistent address. Pronouns simply indicate that their referent is in activation at the time of utterance."*
>
> (Vallduví, 1990: 68, Fn.49)

Successful anaphoric resolution is a matter of descriptive content. DPs can lack enough descriptive content to pick out the intended referent. We also expect that the generalization found in Heim 1982 and Vallduví 1990 can also be violated into the other direction. As we have argued, pronominals DO have some descriptive content, i.e. a restriction on their possible antecedents, namely their phi-features. If this restriction suffices to pick out the intended discourse referent we expect that a pronoun can have access into a set. For testing this prediction we only have to translate Heim's example into German. Here the two NPs under consideration differ in gender:

(72) *Hans hat [[eine Katze]$_x$ und [einen Hund]$_y$]$_z$. [SIE/??sie/die Katze]$_x$ heißt Felix.*
John has a cat and a dog. SHE/she/the cat is called Felix.

There is an interesting issue concerning this example. It looks as if we can maintain a strict version of the restriction that pronouns do not have access into the set. Notice that the second sentence is much better if the pronoun is focussed. It is generally assumed that focussed pronouns can introduce discourse referents (and remember that focus can be taken as a test for being new information). If this conclusion is really valid then we have independent support for our claim concerning Identification. The pronoun can re-introduce a discourse referent. It does so by means of picking it out of a given set. (Therefore the presupposition detected in our instances of Identification arises). Notice that in the examples in (72), where the intended antecedent is contained in a coordinated constituent, we cannot simply say that this discourse referent is not salient or prominent: it has just been activated. In this case we seem to be faced with (structural) inaccessibility. These few remarks should now suffice to show that accessibility interacts with other notions. Furthermore, they are intended to show, how descriptive content on the one hand and salience, prominence and accessibility on the other hand interact. In the next section I will further discuss one particular instance of the above mentioned phenomena, i.e. potential antecedents which occur in a coordinated phrase.

6.5.1. (In)Accessibility and coordination

Assume for the moment that a discourse referent x is most prominent if it is the topic of conversation. Consider under this light the following discourse:

(73) *Peter ging gestern ins Wirtshaus. Dort hat er Hans getroffen. Er blieb jedoch nicht lange.*
'Peter went to the pub yesterday. There he met John. He didn't stay long though.'

The unmarked interpretation for the last sentence is that Peter did not stay very long. The other possibility for anaphorically relating the pronoun *er* to a previously established discourse referent would be linking to *Hans*. Since *Hans* is not the topic of this discourse this is a more marked option. Notice however, that language has a device to pick out a less prominent discourse referent. This involves demonstrative pronominals like *dieser* ('this') or *jener* ('that'). If the pronoun in the last sentence is changed the unmarked option for anaphoric resolution changes, too:

(74) *Peter ging gestern mit seinem Hund ins Wirtshaus. Dort hat er Hans getroffen. Dieser/jener blieb jedoch nicht lange.*
'Peter went to a pub with his dog yesterday. There he met John. This one/that one didn't stay long though.

There is also another option available which makes (somehow metalinguistically) use of the linear ordering:

(75) *Peter ging gestern ins Wirtshaus. Dort hat er Hans getroffen. Ersterer/letzterer blieb jedoch nicht lange.*
'Peter went to a pub yesterday. There he met John. The former/the latter however didn't stay long.'

Here pronouns are used, that make use of the linear ordering of the intended referent: i.e. the *ersterer* or *letzterer*. These pronominals refer to the linear ordering of the relevant antecedents with respect to each other and not to an inherent property of the discourse referent.

A similar distinction between personal pronouns and demonstrative-like pronouns (as the one in (74) can be observed w.r.t. pronominals referring to abstract discourse referents (like propositions, events, facts, etc.). Asher (1993) gives the following example:

(79)

u v y Y U
John (u)
Mary (v)
Acapulco (y)
u took v to y
Z = u + v
U = Z
U had a lousy time

where u,v,y are individuals i.e. the individual (atomic) discourse referents and Y,U correspond to sets of individuals i.e. non-individual (non-atomic) discourse referents. The discourse referent U is induced by the plural pronoun (*they*) whereas Z is created by Summation. Then U and Z are identified. Summation is defined as follows:

(80) (PSUM) Given any finite number of discourse referents $\beta_1,...,\beta_n$ where each β_i may be a plural or singular discourse referent, we may form the *sum* $\beta_1 + \beta_2 + ... + \beta_n$ and use it to replace a *?* in an incomplete condition of a DRS K, if the β_i are all accessible from K. (Asher, 1993: 93)

Consider now an informal representation of the discourse referents in (78) (where irrelevant representation is omitted):

(81)

1)	x
	Peter (x)
2)	y
	Hans (y)
3)	X = (x,y)
	sie (X)
4)	X = (x,y)[37]
	er = ?

Now we can see how Identification comes into play again. Even if we use a definite description, i.e. *der Mann* anaphoric resolution cannot apply successfully, since there

[37] It might be the case that the set created by summation is in fact an ordered set. Thus it might be the case that it looks as follows: X = {x,y}. The ordering relation might again be related to topichood, the topic of conversation (i.e. *Peter*) being the first element. This could explain the fact that if anaphoric resolution takes place the most plausible candidate would be *Peter* (since it is the topic of conversation). Thus we might say that if access into a set is felicitous it can just access the topmost element within the set. However, I still think that this interpretation is a very marked option. Thus I will abstract away from the possibility to assume an ordered set created by summation.

(76) *The "liberation" of the village had been a disaster. [First on a sweep through the town some of the Marines had gone crazy and killed some innocent villagers. To cover up the "mistake", the rest of the squad had torched the village. [To cap it off, the lieutenant called in an air strike.]$_j$]$_i$ At first the battalion commander hadn't believed it/that$_j$.* (Asher, 1993: 226)

According to Asher (1993) stressed *it* may refer deictically but its referent must be the topic of conversation. As indicated if *that* is used it picks up the proposition expressed by the previous clause - the proposition about the *lieutenant's calling in an air strike* (while *this* would pick up the entire topic). There is a difference between the *topic-oriented* use of *this* (and stressed *it*) and the *non-topic oriented* use of *that*. In general demonstrative pronouns are more often used to refer anaphorically to discourse referents that are "further away". Personal pronouns seem to be more "local" pronouns. Consider now the following discourse:

(77) *Peter ist gestern mit seinem Hund ins Wirtshaus gegangen. Dort hat er Hans getroffen. Sie haben Karten gespielt. ??Er hat verloren.*
'Peter went to a pub with his dog yesterday. There he met John. They played cards. ??He has lost.'

Here the pronominal *er* sounds quite odd. This is due to the fact that anaphoric resolution cannot unambiguously determine the relevant antecedent. In the above example two previously established discourse referents have been "grouped together" by means of summation before the pronominal occurs. A set containing two single individuals has been established. In this case the single constituents are not accessible to anaphoric resolution (by the pronoun) anymore. For a more detailed discussion of this fact let me briefly introduce the notion of summation as it is formulated within DRT. Consider the following discourse (taken from K&R 1993):

(78) *Last month John took Mary to Acapulco. Fred and Suzie were already there. The next morning they set off on their sailing trip.*

Here the plural pronoun *they* does not have as single NP as its antecedent. Rather the antecedent has to be constructed out of various parts of the preceding text. Thus we get the following DRS:

are two possible antecedents satisfying the property of being a man. Remember that I have argue that an Identifier can provide further restriction on the set of possible antecedents. The property provided by the identifier should suffice to pick out the relevant discourse referent. So we expect that if an identifier is added then anaphoric resolution can successfully apply. We simply have to provide enough descriptive content to pick the relevant discourse antecedent out of the set of possible referents[38]:

(82) *Peter ist gestern mit seinem Hund ins Wirtshaus gegangen. Dort hat er Hans getroffen. Sie haben Karten gespielt. Der Mann #(der den Hund dabei hatte) hat verloren.*[39]
Peter went to a pub with his dog yesterday. There he met John. They played card. The man #(who had the dog with him) has lost.'

Notice that by means of summation we end up in a situation where a discourse referent is salient (i.e. it has just been mentioned) but it is still not accessible to anaphoric resolution. Consider now again Heim's example discussed above.

(83) *John has a cat and a dog. The cat's/#it's name is Felix.*

In this example the coordinated NP introduces a set of two single individuals. It is impossible to anaphorically refer to either of the NPs within the coordinated phrase by means of a pronoun. The representation for the nominal elements is as follows (irrelevant representation omitted):

(84) x, Y
John (x)
[cat (y) and dog (z)] (Y)

As indicated above both, *cat* and *dog* introduce discourse referents but they are not immediately accessible. The only discourse referent available is Y (but not the single individuals within this set). As opposed to the example discussed above here the

[38] One has to look up the information previously given for the two individuals. Thus we will find one DRS where the condition (with the dog) is associated with x (i.e. *Peter*). Identification can thus license re-introduction of an already introduced (but not accessible) discourse referent (in addition to license 'brand-new' introduction into the discourse).

[39] I think that for the moment we can abstract away from the fact that referring with a definite description to an individual whose name is familiar is not that felicitous. But this is due to other reasons which I am not interested in. However, I think that the point I want to make is clear even in this example, where an intervening factor should be excluded.

definite NP provides enough descriptive content to pick out the intended discourse referent. If *the cat* is used we have enough descriptive content to refer to the actual referent.

(German) pronominals that are linked to sentential constituents are always neuter. We can interpret neuter as 'lack of gender features'[40]. If true then they do not encode any descriptive content, i.e. no restriction on the possible antecedent is provided (except maybe for a feature |-human|). There are many abstract objects (like facts, propositions and eventualities) salient in a given discourse (maybe more than individuals). It is therefore expected that these elements induce the need for Identification more often than other pronominals, which at least have phi features as a restriction. Now let us draw our attention to the behavior of Identification w.r.t. coordinated identifiers. Consider the following sentences:

(85) (a) *Peter hat [es]$_i$ geglaubt, [[daß Maria gerne Bier trinkt] und [daß Hans gerne raucht]]$_i$*
Peter has it believed that Mary likes to drink beer and that John likes to smoke.
(b) *Peter hat [es]$_i$ bedauert, [[daß Maria gerne Bier trinkt] und [daß Hans gerne raucht]]$_i$*
Peter has it regretted that Mary likes to drink beer and that John likes to smoke.

There is a striking difference between the cases discussed above (where coordinated individuals were involve) and the sentences in (85). In case we have coordinated individuals the pronominal which is coreferent to the coordinated NP is plural. In the sentences above, where two CPs (denoting two propositions (85a) or two facts (85b)) are coordinated we find a singular coreferent (cataphoric) pronominal. In fact, it is impossible to have a plural pronominal instead as exemplified below:

(86) (a) **Peter hat sie geglaubt, daß Maria gerne Bier trinkt und daß Hans gerne raucht.*
Peter has them believed that Mary likes to drink beer and that John likes to smoke.
(b) **Peter hat sie bedauert, daß Maria gerne Bier trinkt und daß Hans gerne raucht.*
Peter has them regretted that Mary likes to drink beer and that John likes to smoke.

[40] See also chapter VII. for this assumption.

However, this property of coordination of abstract entities is not only found with 'cataphoric' relations. It can also be observed w.r.t. anaphoric relations. (It is therefore not a particular problem for my analysis). Consider the following discourse:

(87) *Peter hat geglaubt, daß Maria gerne Bier trinkt und daß Hans gerne raucht. Klara hat es/das/*sie auch geglaubt.*
'Peter believed that Mary likes to drink beer and that John likes to smoke. Klara believed it/that/*them too.

Why should this be so? There is an independently motivated explanation for this apparent problem. It has been argued in Asher (1993) that different abstract entities behave differently w.r.t. various operations (like e.g. negation, disjunction, within conditionals, coordination, etc). The question arises whether the obtained entity is still of the same sort before and after one of these operations respectively.[41] We have to ask whether coordination of two abstract objects results in denoting another abstract object of the same sort. Intuitively it seems plausible to assume the following: given a fact F_1 and another fact F_2 then coordination of F_1 and F_2 results in another fact F_3. However, one cannot assume that F_3 is a set that consists of F_1 and F_2. The same is true for propositions. However once F_3 has been introduced than F_1 and F_2 are not accessible anymore.

On the other hand if we try to do the same with individuals it becomes clear that coordinating two individuals I_1 and I_2 does not create another individual I_3. It creates a set consisting of two individuals. This set is not to be analyzed as an individual. It is a different sort of entity. Natural language is sensitive to whether a set is itself treated as a single object or whether this set is still transparent as to whether it contains a number of individuals.

Again pronominals differ from definite descriptions in this respect. As soon as we replace the pronominal with a definite description both options become available: the DP can occur in singular or in plural, no matter whether we are dealing with an anaphoric or a cataphoric relation:

(88) (a) *Peter hat bedauert, daß Maria gerne Bier trinkt und daß Hans gerne raucht. Klara hat die Tatsache/die Tatsachen auch bedauert.*
'Peter regretted that Mary likes to drink beer and that John likes to smoke. Claire has regretted this fact/these facts as well.'

[41] Another question that arises is for example: is a negated event still an event?

(b) *Peter hat die Tatsache/die Tatsachen bedauert, daß Maria gerne Bier trinkt und daß Hans gerne raucht.*
'Peter regretted the fact/the facts that Mary likes to drink beer and that John likes to smoke.

Since events are argued to be less abstract (i.e. they are argued to denote semi-concrete entities in Asher 1993) they should more behave like individuals. Intuitively it seems plausible again that an event E_1 and another event E_2 do not create a third event E_3. Events seem to be unable for closing them under summation. This is true for summation which is not the result of coordination. Thus consider the following example of Asher (1993):

(89) *Fred went shopping for a new pair of sneakers. Mary went to her office to prepare some reports. It took a long time.*

Although matters are not quite simple it seems that the pronominal can only refer to the second event but not to the two events. Events do not sum up that easily. However, if the two events under consideration are somehow connected, then it is again possible to denote a group of events (i.e. to close them under summation). Consider the following minimal pair (again taken from Asher 1993):

(90) (a) *John's hitting of Fred and Mary's getting married were the main events of the week*
(b) *John's hitting of Fred and Fred's subsequent returning of the punches was the main event of the week.*

There is an important issue for the present discussion: once two discourse referents are summed up (by means of coordination) these discourse referents are not available anymore. (cf. Asher (1993) a similar conclusion).

This is an important clue for our analysis. It provides us with independent evidence of the apparently contradictory claim we argued for in case of backwards pronominalization and identification of pronominals. An element which is already established can be (re)-introduced again. There are various possibilities to make an already introduced discourse referent inaccessible for anaphoric resolution. In this case it has to be reintroduced. One of these possibility was the topic of this section, i.e. coordination. Notice that the impossibility for anaphoric resolution to have access into a coordinated phrase is not really surprising. Coordination behaves like an island w.r.t. many phenomena, not just discourse antecedence. For syntactic relations coordinated constituents behave rather opaque (i.e. one cannot extract out of them captured by the

Coordinate Structure Constraint, Ross 1967).

In principle the same remarks hold for other means of creating sets of discourse referents rendering (even salient or prominent) discourse referents inaccessible. We will come back to some of these in the next section. Some of the examples above already showed that definite DPs can (re-)introduce discourse referents in a similar fashion as pronominals. The difference w.r.t. pronominals here depends on the degree of descriptive content. In the next chapter I will explicitly argue for the claim that definite DPs CAN introduce discourse referents.

IV. IDENTIFICATION OF DEFINITE DPs

In this section I will be concerned with definite DPs. On the one hand we will see that the claim that they can introduce (or reintroduce) a discourse referent is independently motivated. To be precise I will argue for the following assumption concerning the interpretation of definite DPs:

(1) the identifyee (the definite nominal) introduces uniqueness
the identifier (the sentential constituent) provides availability in discourse

Notice that this assumption somehow combines two different approaches towards the interpretation of definite determiners. On the one hand it has been claimed that a definite description is a predicate and the proposition arises that this predicate is true of one and only one element. This means that the definite determiner is analyzed as carrying along a uniqueness presupposition (cf. Russell 1905). On the other hand it has been claimed that the definite determiner implies Familiarity in discourse (cf. Heim 1982). In most cases the two properties go together. This is not surprising if one assumes a notion of meaning which is sensitive to discourse. If the definite determiner introduces a unique interpretation of DP, then, in order for this DP to be felicitous, there must be one unique individual available in discourse. This is true for some particular instances of DPs in any case, namely in case the DP refers to an individual which is unique in the world (for example definite descriptions like *the morning star*). Since the model for evaluating the interpretation of a given DP is related to the actual discourse (including some notion of knowledge about the world, or salience by being present in or related to the discourse) the presence of the definite determiner introduces uniqueness w.r.t. this model. Assume that there are several possible discourse referents available (i.e. already established). All of these referents could be denoted by the definite DP *the person*. It is clear that this definite DP does not have enough descriptive content to pick out the actual individual. Since the definite determiner introduces uniqueness it must however be the case that the intended referent refers to a unique discourse referent (out of all the available ones). The identifier then provides enough descriptive content to pick out this referent, i.e. it provides discourse availability (which is reminiscent of Familiarity). Identification is a phenomenon where uniqueness and availability are provided by separate constituents whereas in most instances the two properties occur within one single DP. We will also look at some instances where uniqueness is inherently associated with the meaning of a given element (e.g. *einzig* (only) or superlatives). The interpretation of the identifyee is reminiscent of demonstrative elements: the difference being that availability is made

possible by means of an additional element (the identifier) in syntax whereas in case of demonstratives (or deictic elements) Identification takes place by means of an act of demonstration, which does not need to be accompanied by an overt constituent which makes this discourse referent available.[1]

The claim that a definite determiner can introduce a discourse referent has some implications, which I will address in this chapter. First we will see that in the cases under consideration definite DPs (introducing discourse referents) behave more like 'weak NPs'. This predicts that a syntactic environment which is assumed to just allow (what is referred to as) 'weak NPs' also allows for these kinds of definite DPs.[2] The second crucial implication is that (restrictive) relative clauses as well as CPs which are traditionally referred to as complement clauses pattern alike w.r.t. various phenomena. This is not to say that they are to be analyzed equally. It is just the case that although the relation between a DP construed with a relative clause and a DP construed with a complement clause is syntactically (and semantically) different in various ways they still behave similar w.r.t. Identification. Both, relative clauses and complement clauses are licensed as identifiers (at domain D). This is why I will treat them in a similar fashion. Notice that it is not surprising that a given (primitive) relation can be instantiated syntactically in various different ways resulting in different 'constructions'. As we have said in the beginning: constructions are not primitives but simply an output of the interaction of various different principles.

It is also not surprising that one and the same phenomenon is instantiated by means of different constructions - this is true cross linguistically as well as language internally. Let me briefly say what I think this phenomenon implies for linguistic research (within the generative enterprise) in general. If particular constructions are NOT considered to be primitives then they cannot be universals. For typological research (and by extension also for analyzing a given language) one cannot really investigate a given construction and its various differences in different languages (or in different environments). The line of research should rather try to single out the real primitives of natural languages and then start to look how they are instantiated in various languages. Since all of these primitives interact we expect various different

[1] Notice also that in many cases the identifyee involves a demonstrative like pronominal (or determiner).

[2] The present analysis is also supported by the Persian facts mentioned in section II.2.4. There we have seen that Persian allows for a definite (demonstrative) determiner followed by a noun which is marked with an indefiniteness-suffix. Persian marks uniqueness and Familiarity (or rather Novelty), whereas in German no special Novelty marker is provided.

constructions with different properties as their outputs. I think it is perfectly legitimate to sum up two 'constructions' (like relative clauses and complement clauses) and to look at them from the point of view which claims that they are two (different) instantiations of the same sort of primitive relation. In the course of this section I will therefore abstract away from various different properties the two 'constructions' have, assuming that the differences are the result of other interacting constraints.

1. The representation of Relative Clauses

Let me start with the (simplified and generalized) representation I will assume for the purpose of discussing Identification. Consider first the representation of relative clauses from the point of view of Identification (irrelevant representations omitted):

(2) *Peter hat [die [Frau]$_X$]$_x$ gesehen, [die$_x$ [gerne Bier trinkt]$_Y$]$_y$*
Peter has the woman seen who likes beer drinking

NP: X = set of individuals satisfying the property of being a woman
DP: x = ?
x ⊆ X
CP: Y = set of individuals satisfying the property of enjoying beer drinking
y = female, singular
X ∩ Y = {x,y}
x = y

This representation captures the following properties of Identification. The definite DP comes with 2 indices. One is associated with N (=X) and it denotes the set of individuals satisfying the property of being a woman. Recall that I assume an extensional definition of this set. Notice moreover that there are still two possible denotations involved. N can either denote a set of arbitrary women (though relativized to the domain of discourse). In this case the interpretation of the definite DP corresponds to 'brand-new-introduction' of a discourse referent. The second possibility is that the set denoted by N contains previously established discourse referents (which all have the property of being a woman). In this case the definite DP is assumed to reintroduce an already established discourse referent (similar to what we found in case of Identification of pronominals). The difference to the latter (i.e. Identification of pronominals) is that in case of (definite) lexical nominals X activated by N does not only denote the set of possible discourse referents/antecedents. Rather, we have a further restriction given by N, namely that X must be part of the set of individuals

satisfying the property of being a women.

In a next step the denotation of the entire DP is evaluated. Part of the meaning the definite determiner is uniqueness. Therefore a single individual has to be denoted. However, the set denoted by N can contain more individuals. So at the point of uttering the definite DP the actual discourse referent to be denoted cannot be picked out. That is why we find an incomplete condition (x = ?). The only information one can rely on is that x has to be a subset of X. Here Identification comes into play. The relative clause denotes a property (Y) and (as we will discuss later) by means of containing a relative pronoun it also denotes a discourse referent (y). The denoted property should provide enough descriptive content to pick out the actual discourse referent. In the next step an intersection of the two available sets (X and Y) is created. The result of this intersection is the set (x,y), containing the discourse referent denoted by D and the one denoted by the relative pronoun. If Identification is felicitous there should be no other discourse referent contained in this set, and moreover the two discourse referents should correspond to nominal elements that match in phi-features. If this is so, then subsequently x and y can be equated. Therefore the incomplete condition can be completed, i.e. the identifyee is identified.

It is obvious that I have abbreviated the representation above and for ease of exposition I will continue to do so throughout this chapter. The above representation is simply chosen to capture the facts I am interested in here. I am abstracting away from several facts which are not relevant for the purpose of our discussion. For completeness I will briefly discuss some points which are obviously relevant in general. For doing so let me briefly show how DRT represents relative clauses and how the representation in (2) differs. Consider the following sentence and its (simplified) DRS (drawn from K&R):

(3) (a) *John owns a book which Smith adores* (K&R: 78, (=1.37)

(b)

x y z
Jones (x)
book (y)
Smith (z)
z adores y
x owns y

(K&R: 83, (=1.47)

There is one crucial difference to the DRT representation of relative clauses and the one I am assuming here. In the DRS above the relative clause itself is not associated

with a discourse referent of its own. This is a result of the assumption that relative clauses denote a property not an individual. Standard DRT-treatments of properties do not assume that properties introduce discourse referents. Again I will rely on Asher (1993) in assuming that properties are also associated with a discourse referent. This however concerns the first abbreviation found in (2). Asher (1993) assumes that properties (as well as some abstract objects like propositions) do not introduce real discourse referents. But the DRSs associated with them are assumed to be objects within the model. It is the DRS itself that can act as a discourse referent. However, I think that the representation in (2) is justified because of the following reason. Consider again the representation of the relative clause I am assuming:

(4) *[die [Frau]$_X$]$_i$ [die$_i$ [gerne Bier trinkt]$_Y$]$_x$*
the woman who likes beer drinking

Notice that the representation of the relative involves two indices, Y and y. Y (presumably the index induced by IP) corresponds to a property (again defined extensionally, i.e. the set of individuals who like to drink beer).[3] Notice that this is not the only information the relative clause conveys. I assume that the relative pronoun is also associated with an index. In a way the representation of the relative clause corresponds to the representation of nominals (where NP is associated with a property interpretation and the DP is associated with the actual discourse referent). I furthermore assume that the index of the relative pronoun (y) in Spec CP percolates to CP itself (either by means of direct percolation from SpecCP or by means of Spec-Head-Agreement with C^0 and subsequent percolation from the head position to CP). The relative pronoun imposes some restrictions on the interpretation of the relative clause by virtue of its phi-features. Since x is identified with y (after intersection of X and Y) I will continue to represent relatives as follows:

(5) *[die [Frau]$_x$]$_i$ [die$_i$ [gerne Bier trinkt]]$_x$*

Due to Identification the DP and the relative clause behave as if they were coindexed (which is mainly a result of the presence of the relative pronoun). Therefore the Novelty Condition can apply resulting in the desired precedence constraint: the identifyee (i.e. the nominal) must precede the identifier (the relative clause).

[3] Most likely a relative clause is interpreted as a property because by means of wh-movement of the relative pronoun it contains an open place, i.e. a variable.

1.1. A crosslinguistic consequence

This representation has an interesting crosslinguistic prediction. It is the result of two claims I make. The first assumption is that the obligatory rightness of a relative clause is derived by the Novelty Condition. The Novelty Condition simply says that a discourse referent introducing DP cannot be preceded by another element bearing the same index. This assumption seems dubious once we consider the fact that there are languages which have head-final relative clauses. However, constructions are just the output of various constraints. And here the second claim I am making becomes relevant. This second assumption my analysis relies on is that coindexing between the relative clause and the head-NP is due to the presence of a relative pronoun. Therefore the following crosslinguistic prediction arises: We expect that languages which exhibit so called head-final (or even head-internal) relative clauses do not have relative pronouns. In this case the relative clause is not associated with a discourse referent y (but only with the discourse referent associated with the Y, i.e. a property, which is subject to different restrictions). Interestingly the same prediction (though on basis of different assumptions) is made in Kayne (1993). He refers to Downing (1978) for arguing that this prediction is apparently correct. Consider the following quotation of Downing (1978):

> *"Most strikingly there are no relative pronouns in prenominal clauses [...]"*
> (Downing, 1978: 392)

Notice that I will argue empirically against Kayne's (1993) approach for relative clauses in Appendix II. Thus if it is true that his analysis is not on the right track then my way to derive this universal is superior. For the moment let us assume that the prediction is correct.[4]

[4] Moreover I predict that languages which allow for preposing a relative clause cannot do so anymore if the head NP introduces a new discourse referent. This is a correct prediction judging from Lehmann (1979):
"All types of relative clauses except for the preposed one can in this way be a constituent of a Rheme.[...] Although a relative clause in Hindi is in principle 'movable' it cannot be preposed:
us-ne ek jhīl dekhi, jo bahut barī thī
3-INST a lake saw [$_{REL}$ NOM SG very big was]
He saw a lake which was very big. (Lehmann: 420f., translation and emphasis MW)
According to Lehmann the impossibility for the relative clause in this example follows from the fact that the syntactic environment forces a Rheme interpretation of the NP. In our terms this can be interpreted as to saying that the relative clause can be preposed but only if the NP does not introduce a new discourse referent. This however implies that in German relative clauses are just licensed if the head NP (re-)introduces a new discourse referent.

2. The representation of 'complement clauses'

Let me now turn to complement clauses and the representation I will use. Consider the following sentence:

(6) *Peter hat [die [Tatsache]$_X$]$_y$ erwähnt, [daß Maria gerne Bier trinkt]$_x$.*
Peter has the argument presented that Mary likes beer drinking

Matters are more complicated here and we cannot make use of the same representation as assumed for relative clauses. Consider how a similar representation would look like:

(7) *Peter hat [die [Tatsache]$_X$]$_y$ erwähnt, [daß [Maria gerne Bier trinkt]$_Y$]$_x$.*

NP: X = set of entities satisfying the property of being a fact
DP: x = ?
$x \subseteq X$
IP: Y = set of entities satisfying the property of 'Mary likes drinking beer'
CP: y = the abstract entity 'that Mary likes drinking beer
$X \cap Y = (x,y)$
x = y
[thus y = a fact]

The problem arising here is twofold. First notice that it is unclear what the relation between Y and y is, i.e. how the instantiation relation between y and Y is established. This has to do with the fact that IP here does not contain a variable (as it does in case of relative clauses). It is a saturated argument and should not have a property interpretation. I will however not go into this problem, which is beyond the scope of the present discussion.[5] The following remarks should suffice to justify the representation as it stands, i.e. the internal representation of the clause is not of any interest here since the only thing which is relevant for the present analysis is the possibility to coindex NP and CP.

[5] However notice that the assumptions I have made so far would lead to an interesting line of research concerning the interpretation of sentential elements. If it is true that sentential elements are also associated with a discourse referent (and I have argued that this is a necessary assumption) then we might expect a correspondence law as it is assumed for nominals (following V&Z). Sentential elements might be associated with a type and a token interpretation in the sense of V&Z.

There is still another problem induced by the representation above indicated by the information *[thus y = a fact]*. Simply creating an intersection between X and Y would not give us the set of (x,y) since there is nothing inherent to the CP that would tell us that the CP denotes a fact. It might as well be a proposition or a possibility or an eventuality. It is the content of the DP itself which specifies that the CP denotes a fact. Notice however, that this is a general property of *that-clauses*. Their exact interpretation is dependent on various interacting properties, e.g. the N or V they are dependent on, the position they occur in[6] etc. They have some sort of chameleon-like property (similar to bare plurals or indefinites whose interpretation depends on various factors). Moreover, it is not a standard assumption that CPs are associated with a discourse referent. Again I will assume (with Asher 1993) that they behave like they would by virtue of being associated with a DRS which counts as an object in the model. Again this can be seen by means of the possibility to anaphorically refer to such abstract entities as exemplified below.

(8) *Maria hat bedauert, [daß Hans kein Bier trinkt]. Klara hat dasldiese Tatsache jedoch gutgeheißen.*
Mary has regretted [that John no beer drinks]. Claire has that/this fact though approved.

Technically this can be implemented by saying that a DP takes an individual (or a discourse representation) as its argument whereas a CP takes a DRS as its argument. This amounts to saying that the DRS associated with the CP acts as the discourse referent. The second evidence for the constant-like behavior of CPs (taken from Asher 1993) is their behavior w.r.t. the following quantificational pattern of (valid) reasoning:

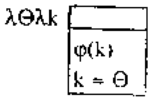
(9) (a) *John believes everything that Mary believes.*
(b) *Mary believes that Susan is smart*
(c) *So John believes that Susan is smart.*

[6] Consider the well-known difference in interpretation between the following sentences:
(i) *That the president was dead was reported by the UPI*
(ii) *That the president was dead was not reported by the UPI*
(iii) *The UPI reported that the president was dead*
(iv) *The UPI didn't report that the president was dead*
If the CP is in subject position then the proposition it denotes is presupposed to be true (it is a fact). It is not influenced by negation of the matrix predicate. On the other hand if the CP occurs in object position no such presupposition arises and negation of the matrix predicate has an influence.

In the first sentence we quantify over the structures induced by the *that* clause. In the second premise the structure introduced by the *that* clause serves as an instance of the antecedent of the first premise. If one interprets such quantification referentially then it can be assumed that DRSs are objects in the domain of a DRT model (and quantifiers like *everything* may range over them).

The problem that there is nothing inherent to the CP specifying that it denotes a fact (and therefore we cannot arrive at the correct interpretation simply by means of creating the intersection of X and Y) needs some more attention. In accounting for this problem I will again rely on Asher (1993), who provides an analysis for these complex NPs. I will present his representation in order to justify the abbreviated version where we simply coindex DP and CP. For the sentence in (10) Asher gives the following representation:

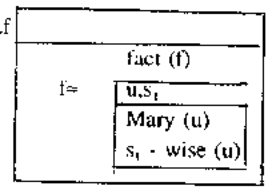
(10) *The fact that Mary is wise reflects well on her teachers.*



(Asher 1993: 174ff)

where k is an abstract entity discourse referent, while Θ ranges over DR-theoretic structures of abstract discourse referents. He assumes that the head noun of this complex NP determines the interpretation of the discourse referent. Thus the noun *fact* dictates that the discourse referent argument must be of fact-type. This is encoded in the following representation: f is a fact discourse referent variable that is shorthand for type information encoded as condition on the predicative DRS:

(11)



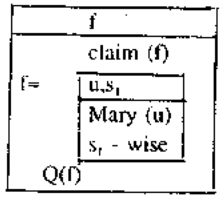
(Asher, 1993: 176)

The predicative DRS then combines with the translation of the determiner *the*.[7] The

[7] Notice that in case of an (abstract) noun like *fact* there is a difference between identification by means of a complement clause or a relative clause: the former instance is just legitimate with the definite determiner whereas in the latter case we can have either a definite or an indefinite determiner:
(i) *Die/*eine Tatsache daß Peter gern Bier trinkt kann nicht geleugnet werden.*
the/*a fact that Peter likes beer-drinking can not denied be.

result is the following partial DRS. The discourse referent introduced by the determiner inherits the type restrictions in the predicative DRS:

(12) λQ



(Asher, 1993: 176)

What interests us here is the fact, that this sort of representation reflects the claim that the denotation of the CP IS equated with the denotation of the DP. It is assumed that the DP introduces a *characterization condition* on the content of the whole CP. Therefore I take it that we can legitimately stick to the simplified analysis where the CP and the DP are coindexed - abstracting away from all the problems induced by reference to abstract objects in general.

Additional support for coindexation of DP and CP stems from the fact that the relevant DPs and CPs can be identified 'overtly' by means of the copula (which I take to be a construction instantiating Identification). The following sentence is an example:

(13) *[Die Tatsache]$_x$, von der wir reden, ist [daß Maria gerne Bier trinkt]$_y$.*
'[The fact]$_x$ we are talking about is [that Mary likes to drink beer]$_y$.
x = y

The analysis for complex NPs denoting abstract objects given in Asher (1993) also justifies the claim that relative clauses and 'complement' clauses are treated in a similar fashion. Consider the following quotation of Asher (1993):

> *"The very same that-clause may characterize a wide variety of abstract entities - thoughts, propositions, facts, possibilities [...] these different NPs give rise to different and incompatible discourse referents. that-clauses put conditions on these discourse referents."* (Asher, 1993: 172)

'The/*a fact that Peter likes to drink beer cannot be denied.'

(ii) *Daß Peter Bier trinkt, ist die/eine Tatsache, die nicht zu leugnen ist.*
that Peter beer drinks, is the/a fact, which not to deny is
'That Peter drinks beer, is the/a fact, which cannot be denied.'

I have nothing of interest to say about this minimal pair.

What is important is the claim that clauses put conditions on discourse referents. Conditions on discourse referents can be of various sorts. In most of the cases conditions are simply predicates denoting properties. Relative clauses also denote properties, i.e. conditions on the discourse referent. This is what complement clauses and relative clauses have in common: they put conditions on the discourse referent - in our terms they convey information to identify the actual discourse referent. We can now turn to the predictions and implications the present analysis suggests.

3. Definite DPs introducing discourse referents

The present section has two main goals. In the first part I will argue that definite NPs are not always subject to the Familiarity Condition. There are phenomena (independent of Identification) which show that definite NPs can introduce a discourse referent.[8] This provides independent evidence for the necessary claim of our analysis that those instances of definite DPs we are interested in (i.e. if they act as identifyees) really CAN introduce or re-introduce a discourse referent. The second part provides evidence that in case of Identification of a definite DP it really introduces a new discourse referent. For doing so we will use tests that involve environments that are independently argued to be environments where just elements conveying new information are legitimate. An important consequence of this section is that the definite determiner cannot be equated with 'strong NPs' (or old information). Rather it comes with a uniqueness presupposition. The availability in discourse is provided by the identifier. In case of re-introducing the discourse referent availability is related to Familiarity whereas in cases of 'brand-new' introduction of a discourse referent by a definite DP the identifier *introduces* availability. I will provide evidence that the identifier licenses a definite DP to introduce (or re-introduce) a new discourse referent.

3.1. The obligatoriness of a definite determiner

The first piece of evidence that definites can introduce a discourse referent stems from the fact that certain NPs are (syntactically) only wellformed with the definite determiner no matter whether the referent has already been introduced or not. The

[8] As already mentioned in the chapter III. Heim (1982) also mentions various instances of definite DPs which seem to violate the Familiarity Condition, i.e. they seem to be novel. The relevant mechanism for doing so is 'accommodation', which states that cross-referential linking to some already present file card can license definite DPs to introduce a new file-card. (I refer the reader to Heim (1982) as well as Lewis (1979) for a discussion of accommodation).

obligatoriness of the definite determiner can be seen even with sentences that are uttered out of the blue:

(14) *Peter hat den/*einen besten Biertrinker getroffen*
Peter has the/*a best beerdrinker met
'Peter has met the/*a best beerdrinker.'

(15) *Peter hat den/einen einzigen Biertrinker getroffen*[9]
Peter has the/*an only beerdrinker met
'Peter has met the/*an only beerdrinker'

(16) *Peter hat den/*einen third Mann gesehen.*
Peter has the/*a third man seen
'Peter has seen the/*a third man.'

These sentences show that the definite determiner must be able to introduce a discourse referent. Otherwise it would not be possible to talk about *'the best N'* or *'the only N'* etc. The relevant DP could never be introduced in the discourse. Notice that this phenomenon is again an instantiation of the optionality hypothesis. If syntactically there is no option, i.e. the definite determiner is syntactically necessary, then it is not necessarily related with its usual discourse function: it does not have to be familiar in the discourse. It is also interesting to notice that all the examples above involve a modifying element which is related to uniqueness. In the light of what we have said so far it is not unexpected that the definite determiner is necessary: in all the examples above uniqueness of the discourse referent corresponds to availability, since there is just one particular individual (relevant for the present discourse) that fulfills the property denoted by the modifying element. In these cases the modifiers (superlative, only, ordinal number) provide a sufficient property to single out the relevant discourse referent. There need not be an indefinite DP preceding the definite DP to make the discourse referent available. A similar conclusion can be drawn from the following example, where it is also the case that the definite determiner is the only option (similar examples are discussed in Vergnaud 1974):

(17) *Peter hat das/*ein Paris seiner Jugend wiederentdeckt*
Peter has the/*a Paris his$_{GEN}$ youth rediscovered

As I briefly sketched in chapter II.-III. I assume that proper names are also associated with two indices: X and x. Assume that a Name basegenerated in N obligatorily has

[9] The sentence in (15) has an interpretation where the indefinite determiner is wellformed (it corresponds to *single*). Still the definite determiner does not imply Familiarity.

to move to D° resulting in the interpretation that x = X. Since x is associated with uniqueness it follows that if there is just one member of this set it can move there. However if a definite determiner is added the set-interpretation is still available. The above example is interpreted as if there were more possible entities satisfying the property of the predicate (i.e. different 'kinds' of *Paris*).

In these cases the definite determiner does not play the same role as in other contexts. In the cases above it can introduce a discourse referent because the NPs under consideration are associated with uniqueness. This means that the a unique property is attributed to the intended discourse referent. (Notice this is reminiscent of Donellan's notion of *attributive* reading. Informally, it doesn't matter whether there is an actual discourse referent present in the discourse or not. the property ascribed to the potential individual would suffice in any case to single it out. There is empirical evidence that this is indeed the case.

Consider copula-sentences, which are ambiguous between a predicative and an identificational reading. The former involving an NP which is not referential the NP being predicated of the subject. Here the copula is present for purely syntactic reasons. On the other hand for the identificational reading it is assumed that the copula acts like a transitive verb. The NP involved is supposed to be *referential*. The difference can easily be seen if the predication-relation is negated.

(18) (a) *Gustav ist ein Mann*
Gustav is a man

(b) *Gustav ist der Mann mit dem Hut*
Gustav is the man with the hat

(19) (a) *Es ist nicht der Fall daß Gustav ein Mann ist*
It is not the case that Gustav a man is
'It is not the case that Gustav is a man'

(b) *Es ist nicht der Fall, daß Gustav der Mann mit dem Hut ist*
It is not the case, that Gustav the man with the hat
'It is not the case that Gustav is the man with the hat.'

In case of negation of the predicative sentence (19a) we do not talk about a man who Gustav is not, we just deny the fact that Gustav is an individual with the property of being a man. If however the identificational context is negated (19b) a different interpretation arises. It is still presupposed that *the man with the hat* exists. Here Identification itself is negated. (19b) asserts that *it is not Gustav who the man with the hat is* (but somebody else). Definite NPs are in general assumed to force the

identificational interpretation. However if any of the definite DPs (which are just wellformed with the definite determiner) is used the situation changes. In that case the copula sentence shows the same kind of ambiguity as discussed above:

(20) (a) *Es ist nicht der Fall, daß Gustav der einzige Mann ist.*
'It is not the case, that Gustav is the only man'
(b) *Es ist nicht der Fall, daß Gustav der schönste Mann ist.*
'It is not the case that Gustav is the most beautiful man'

In both of these examples we do not necessarily imply that *the only man exists* or that *the most beautiful man exists and Gustav is not this particular man*. The sentences are ambiguous between the presuppositional and the non-presuppositional reading.[10]

The second context where the same conclusion can be drawn involves another predicational context which does not involve a copula:

(21) (a) *Ich halte Peter für einen Idioten*
'I consider Peter an idiot'
(b) *Ich halte Peter für den Idioten*
'I consider Peter the idiot'

In case of an indefinite NP the sentence means that the speaker thinks of Peter as having the property of being an idiot. Again, if we look at the same construction involving a definite determiner it is implied that *there is a particular idiot* and the speaker thinks that *Peter is this idiot*. Thus the existence of a particular idiot is presupposed and Peter is identified with this idiot. Again the definite DPs under discussion here differ w.r.t. to this predicational relation. Both readings are available:

(22) (a) *Ich halte Peter für den einzigen Idioten*
'I consider Peter the only idiot'
(b) *Ich halte Peter für den größten Idioten*
'I consider Peter the worst idiot'

What one could conclude from the (well-known) generalization, that the presence of the definite determiner does not imply referentiality is that those particular DPs are not associated with a discourse referent, similar to what standard DRT-treatment claims to

[10] Notice again that this ambiguity is in accordance with the Optionality Hypothesis introduced in section III.3.: once no syntactic option is possible then there is an option in interpretation.

be true for (indefinite) predicative NPs. However, I will argue in chapter V. that even predicative NPs are associated with a discourse referent (of a special kind). What is important here is that for predicative (indefinite) NPs on the one hand and attributively used (definite) NPs on the other the presence of the (in)definite determiner does not imply anything as to whether the discourse referent is introduced or already established. As we said it is associated with the meaning of the relevant modifying element that there is just one unique discourse referent. Since it is unique it is at the same time available. Thus the determiner is not enough for evaluating the Novelty-vs. Familiarity Condition - since it solely depends on the nominal element which determiner is licensed. However, it is argued here that both 'predicative' as well as 'attributive' NPs CAN introduce a discourse referent. And I will therefore assume that if they introduce a new discourse referent they also obey the Novelty Condition (and therefore the Precedence Constraint for Identification can still be maintained).

3.2. The identifier licenses a definite DP to introduce a discourse referent.

Let me now turn to the second prediction. There is important evidence showing that definite DPs can introduce discourse referents. Identifiers contribute to the interpretation i.e. as we have said they provide enough content to make the relevant discourse referent available. What the definite determiner does here is not indicating Familiarity, but solely uniqueness. N (associated with X) either denotes a set of already established discourse referents or it introduces a set of possible referents (all having the same property). The identifier then either ascribes an already familiar property, i.e. it provides availability in order to re-introduce a discourse referent. Alternatively it can also introduce availability by means of providing a property which the speaker takes to be a property of only one of the possible discourse referents.[11] First let us look at

[11] Notice that *uniqueness* as the relevant interpretation associated with the definite determiner has also been argued for in Kadmon (1990) within a DRT framework. There the following definition is given:
B_K: = {X:∃K' accessible from K s.t. K'≠ K and X∈ $U_{K'}$}
The Uniqueness Condition
Let α be a definite NP associated with a variable Y, let $K_{\alpha c}$ be the local DRS of α, and let K be the highest DRS s.t. K is accessible from $K_{\alpha c}$ and Y∈ U_K.
α is used felicitously only if for every model M, for all embedding functions f, g verifying K relative to M, if ∀X ∈ B_K f(X) = g(X) then f(Y) = g(Y). (Kadmon 1990: 293)
According to this definition uniqueness is relativized w.r.t. to an accessible domain. He furthermore states that there is no particular constraint imposed on the uniquely identifying property. It can be any property P present in the DRS. What I am arguing here is simply that this uniquely identifying property can be provided in the sentence, just in case the definite NP itself does not provide such a property.

a test already established in section II.2. i.e. question answer pairs. The paradigm below shows that it is exactly in contexts where the DP is identified that a definite determiner is licensed to introduce a discourse referent. This generalization holds both, for relative clauses (23) as well as complement clauses (24):

(23) Q: *Weißt du, ob Maria ein Buch gelesen hat?*
'Do you know whether Mary has read a book? '
A1: *Ja, Maria hat ein/#das Buch gelesen.*
'Yes Mary has read a/#the book.'
A2: *Ja, Maria hat ein/das Buch gelesen, das Hans ihr geschenkt hat*[12]
'Yes, Mary has read a/the book John gave to her.'

(24) Q: *Weißt du ob Hans ein Argument präsentiert hat?*
'Do you know whether John presented an argument? '
A1: *Ja, Hans hat ein/#das Argument präsentiert*
'Yes, John presented an/#the argument'
A2: *Ja, Hans hat das Argument präsentiert, daß Maria gerne Bier trinkt*
'Yes, John has presented the argument, that Mary likes to drink beer'

If a question involves an indefinite DP it is not the case that this indefinite DP already introduces a discourse referent[13]. The answer for a question like (23) or (24) can only involve an indefinite NP. However, if a relative clause or a 'complement clause' is added (i.e. if the DP is identified) a definite DP can be uttered and still the result is a felicitous answer. Here it cannot be denied that the discourse referent was not introduced before the occurrence of this definite DP. I take these examples as crucial evidence for the claim that definite DPs can introduce a discourse referent if they are identified. This phenomenon (however, only concerning relative clauses) is also mentioned in Jackendoff (1977):

> *"A relative clause may be definitizing i.e. render an otherwise indefinite NP definite, just in case it provides grounds for unique identification which were not present in the NP without the relative."* (Jackendoff, 1977: 196)

[12] According to Fn.11 we expect that it is also possible to have a definite DP without an Identifier in case the NP itself provides enough descriptive content. Consider the following example (pointed out to me by Edwin Williams, p.c.):
(i) Q: *Weißt du, ob Maria ein Buch gelesen hat?*
'Do you know, whether Mary has read a book?'
A: *Ja, Maria hat den neuen Roman gelesen*
'Yes, Mary has read the new novel.'

[13] cf. the discussion in section III.2. where the same phenomenon was observed.

The examples in (23-24) show an important assumption of my analysis (concerning the relevance of domain D). The relevant contrast appears only in case the sentences are considered w.r.t. question-answer pairs. For pure syntactic wellformedness the contrast does not show up. (Remember that the same effect was true for Identification of pronominals discussed in chapter III.). The sentence given as the first answers respectively do not show a contrast w.r.t. the determiner if we simply look at them when uttered out of the blue.

Notice that there are syntactic environments that are assumed to only allow for special kinds of DPs. This observation is known as the *definiteness* (or more recently the *specificity*) *effect*. There are two competing approaches two capture this effect. One is to ascribe the relevant property to a purely syntactic feature (i.e. to a syntactic property of the relevant) determiner. The other approach is more pragmatic in nature, claiming that the relevant distinction is to be drawn in terms of new or old information, respectively. The present discussion provides evidence that at least for some instances of the *definiteness effect* the latter approach is on the right track. The line of argumentation is as follows. In this chapter I have argued that identifiers can license a definite DPs to introduce a new discourse referent (by providing enough descriptive content). For these instances of definites the two competing analyses for the *definiteness effect* make different predictions. The syntactic approach predicts that definite DPs are always excluded from environments which do not allow for definite DPs, no matter whether it introduces a new discourse referent or not. On the other hand the pragmatic approach predicts that these definite DPs are well-formed in these environments (which are interpreted as environments that only allow for new information). Notice that in these cases we do not have to provide a special context, rather well-formedness can be judged on basis of isolated sentences. The following sections provide evidence that at least some instances of the *definiteness effect* should be reinterpreted as an instance of a *Familiarity effect*. Moreover it provides evidence that syntactically identified definite DPs behave differently than simple definite DPs.

3.2.1. Presentational Contexts

First let us look at examples referred to as *presentational sentences* (this term already suggests the line of argumentation we pursue here - presentational can be equated with 'introducing a new discourse referent'). In German this is the case in sentences where we find the so called *Vorfeld-es* (i.e. an expletive in SpecCP):

(25) (a) *Es kam ein Mann zur Tür herein.*
It came a man to-the door in.
(b) *#Es kam der Mann zur Tür herein.*
It came the man to-the door in

(26) (a) *Es wurde ein Argument präsentiert*
It was an argument presented
(b) *#Es wurde das Argument präsentiert*
It was the argument presented

It can be observed that the subject of a V2 sentence cannot be definite if the *Vorfeld-es* occupies SpecCP. However, if the definite DP is identified (and thus licensed to introduce a discourse referent) the situation changes. In this case a definite DP is perfectly wellformed within a presentational environment:

(27) *Es kam der Mann zur Tür herein, den Maria schon lange kennenlernen wollte.*
It came the man to door in, who Mary already long getting-to-know wanted.
'The man, that Mary always wanted to get to know, entered the room.'

(28) *Es wurde gestern das Argument präsentiert, daß die Erde rund ist.*[14]
It was yesterday the argument presented that the earth round is.

A similar conclusion has been drawn for English *there-insertion sentences* (TIS). There have been many attempts to explain the constraints imposed on the nature of the postverbal nominal (cf. Ward & Birner 1994 for a recent discussion of the various approaches, which involve syntactic, semantic as well as pragmatic explanations). Although it is the most common approach to claim that there is a *definiteness* or a *specificity* effect involved, there is a variety of examples (mostly involving definites DPs) that contradict this generalization. For the present discussion it is noteworthy that it has been argued in Prince (1992) that the relevant generalization is as follows. The

[14] The same effect has been observed in Guerón (1980), though analyzed slighly different. I will discuss her account in Appendix I. Suffice it to say that she concludes that certain definite DPs can be interpreted as quantificational (i.e. attributive). Guerón gives the following examples:

(i) *Es sind schon 5 Gäste betrunken*
There are already 5 guests drunk

(ii) **Es sind schon die Gäste betrunken*
There are already the guests drunk

(iii) *Es sind schon die Gäste betrunken, die immer betrunken sind.*
There are already the guests drunk, who always drunk are

It is also interesting that Guerón observes that the possibility to extrapose is also influenced by the predicate involved. It is basically 'presentational' verbs which allow for extraposition (in English). This suggests that in case of other verbs, the DP cannot really be novel. Then the relevant relative clause does not act as an identifier, rather it might be reinterpreted as a nonrestrictive relative clause.

postverbal nominal in TIS represents information that the speaker assumes to be not *shared knowledge*. Those nominals are argued to be *hearer-new* information. Ward & Birner (1994) discover 5 different types of hearer-new referents:

I. *Hearer-old entities marked as hearer new*
II. *Hearer-new tokens of hearer old types*
III. *Hearer-old entities newly instantiating a variable*
IV. *Hearer-new entities with uniquely identifying descriptions*
V. *False definites* (Ward&Birner, 1994: 171)

The relevant generalization is again that the postverbal-nominal has to convey (some sort of) new information. The classification in I.-V. above is relevant for the present discussion. Some of them instantiate the phenomenon we are considering: i.e. a definite DP can (re-)introduce an already established discourse referent. Similar to Identification of pronominals we find that the definite NP can at the same time contain old AND new information. Let me just pick out one particular set of examples, namely class IV. These examples are especially relevant for the present discussion since they are related to the uniqueness presupposition. I have argued that a definite DP denoting a discourse referent that is unambiguously identifiable can introduce a new discourse referent. Ward & Birner (1994) give the following examples

(29) *There was the tallest boy in my history class at the party last night*
(30) *There are the following reasons for this bizarre effect...*
(31) *...and then there is this line which comes out to the outer glide slope aim point*

Notice that the examples above involve a superlative (29) cataphoric reference (30) and deictics (31). All of these properties have been argued to be crucial for Identification at domain D. If uniqueness is an inherent property of the DP, then Identification is not necessary in order to pick out the relevant referent. This is true for superlatives. (30) shows that cataphoric relations are not only possible sentence internally (instantiated as Identification) but also intersententially. The adjective *following* indicates that the relevant property of the discourse referent introduced still needs to be provided. (31) involves a DP containing a demonstrative determiner, again indicating that the relevant property of the intended discourse referent will be provided. In all the examples above definite DPs can introduce a new discourse referent but solely because availability in discourse is introduced as well. This means that syntactic environments are really sensitive to discourse properties and not only to syntax inherent features (like definiteness). I take this to be indirect support that S-structure has to be relate to an interface level which encodes discourse properties, i.e. domain D: Identification

relevant for domain D (i.e. providing enough descriptive content for a discourse referent to being picked out) can render a definite DP possible in environments where its occurrence is usually not legitimate.

3.2.2. Other contexts

In order to look at some other contexts relevant for the present discussion let me briefly introduce a generalization given in de Jong (1987). Here I am not concerned with the adequacy of her analysis. It simply provides environments where we can test the prediction that identified DPs behave differently (i.e. they behave on par with 'weak DPs' because they are licensed to introduce a new discourse referent). De Jong (1987) suggests that the feature definiteness vs. indefiniteness should not be considered as a dichotomy, rather she suggests two features (F_1 and F_2) that can be specified [+] and [-], respectively. The two features induce a four-way-distinction among possible determiners.[15] $[+F_1]$ determiners correspond to those that are blocked in TIS whereas $[+F_2]$ determiners are the ones which are allowed in partitives. Examples are given below:

(32) (a) *There are *the/*all/*most/three boys in the garden*
(b) *Some of the/*most/*twenty boys are red-haired.*[16] (de Jong 1987: 271)

de Jong gives a classification of English (and Dutch) determiners w.r.t. to the two features suggested. The definite determiner *the* is classified as $[+F_1]$ and $[+F_2]$. In addition she assumes the following filters on the occurrence of various determiners where again the features she suggests seem to be relevant:

[15] For completeness let me give the definition of the relevant features:
A determiner D is
- *$[+F_1]$ iff the definition of D_tAB presupposes A to be a nonempty set*
- *$[-F_1]$ otherwise*

A determiner D is
- *$[+F_2]$ iff D is antisymmetric*
- *$[-F2]$ otherwise*

A determiner D is antisymmetric iff for every model in which D_tAB & D_tBA is true, A = B is true as well. (de Jong, 1987: 277)

[16] What is crucial for her argument is that determiners like *most* cannot occur in both contexts, whereas others behave complementary. She therefore concludes that a simple binary distinction does not suffice.

(33) (a) *$[+F_1]$ years ago, *$[+F_1]$ meetings later, *$[+F_1]$ summers ago, *$[+F_1]$streets ahead
(b) The path is *$[+F_1]$ meeters long, *the box weighs *$[+F_1]$ ton
(c) it took *$[+F_1]$ seconds
(d) I have *$[+F_1]$ brothers
(e) a house with *$[+F_1]$ windows, a girl with *$[+F_1]$ brothers
(f) *$[+F_1]$ questions came into his mind
(g) She draws *$[+F_1]$ conclusions
(h) *$[-F_2]$ following books are sold out (de Jong 1987: 274)

The filters suggested by de Jong (1987) seem to be at play in German also, as exemplified by the following examples:

(34) (a) *Die Demonstration ist *die/*alle/*beide/zwei/viele/wenige Jahre her.*
'The demonstration is *the/*all/*both/two/many/few years ago'
**Die/zwei Treffen später wurde der Antrag doch genehmigt.*
*The/two meetings later was the application yet granted
*Maria hat Peter vor *den/vielen Sommern kennengelernt.*
Mary has Peter before *the/many summers met
*Das Fischgeschäft befindet sich nur *die/wenige Straßen entfernt von hier.*
The fishshop finds itself only *the/few streets away of here.
(b) *Der Weg ist *die/zweihundert Meter lang.*
The way is *the/two-hundred meter long
*Die Kiste wiegt *die/dreißig Kilogramm*
The box ways *the/thirty kilograms
(c) *Wir benötigten *die/30 Sekunden.*
We need (*the) 30 seconds
(d) *Peter hat *die/viele Brüder*
Peter has *the/many books
(e) *ein Haus mit *den/vielen Fenstern*
a house with *the/many windows
(f) **Die/zwei Fragen kommen mir in den Sinn*
*The/two questions come myself in the mind
(g) *Maria zieht *den/keinen Schluß aus diesem Faktum*
Mary draws *the/no conclusion from this fact

In all these examples the definite determiner is odd, whereas a $[-F_2]$ determiner is possible. Wellformedness is syntactically determined. We do not need a special context to get the contrast. The situation (in many of the examples) changes if we add an

identifier. It can be observed that the identifier licenses a definite determiner in a context where it is not legitimate otherwise:

(35) (a) *Genau die zwei Treffen später wurde der Antrag doch genehmigt, die Peter versäumt hat.*
Exactly the two meetings later was the application yet granted, which Peter missed has.
Maria hat Peter vor DEN Sommern kennengelernt, die das Land in die Trockenheit stürzte.
Mary has Peter before THE summers met, which pushed the country into the aridity
Das Fischgeschäft befindet sich genau die (zwei) Straßen entfernt von hier, die gestern überschwemmt wurden.
The fishshop finds itself exactly the two streets away of here, which yesterday flooded were

(b) *Der Weg ist DIE zweihundert Meter lang, die Peter gestern vermessen hat.*
The way is THE two-hundred meters long, which Peter yesterday measured had
Die Kiste wiegt (genau) DIE dreißig Kilogramm die noch erlaubt sind.
The box weighs (exactly) THE thirty kilograms, which still allowed are

(c) *Wir benötigten genau DIE 30 Sekunden, die noch erlaubt waren.*
We needed exactly THE 30 seconds, which still allowed were

(d) *Jeder hat DIE Brüder, die er verdient.*
Everybody has THE brothers, who he deserves

(e) *Ich sah ein Haus mit den Fenstern, die Hans erfunden hat.*
I saw a house with the windows, which John invented has

(f) *Hier kommt mir die Frage in den Sinn, die Maria schon gestellt hatlob Maria wirklich eine Spionin ist.*
Here comes myself the question in the mind, which Mary already asked has/whether Mary really a spy is.

(g) *Maria zieht aus diesem Faktum den Schluß, den jeder ziehen würdeldaß sie gehen sollte.*
Mary draws from this fact the conclusion, which everybody draw wood/that she leave should

If my analysis is on the right track this suggests that the relevant feature is again Novelty rather than something else. However, it is not the purpose of the present thesis to give an account for the definiteness effect. I leave the question of the exact nature of definiteness or specificity as a matter of future research.

4. Differences between Identification of pronominals and Identification of lexical NPs

We have already seen in chapter III. that pronouns and definite descriptions differ in various ways. There is one obvious difference between full lexical DPs and pronominals. This property concerns the possibility for the head N of lexical DPs to introduce a set itself (i.e. the set of individuals satisfying the property denoted by N). There need not be a previously established set - as it is necessarily the case for pronominals (since their X denotes the set of possible i.e. already established discourse referents). Therefore we expect different effects concerning the presupposition we arrive at. We have seen in section III.2. that the identifier cannot be associated with new information, thus no focus is possible within the sentence nor is it possible to associate a focus particle with the identifier. For the case of definite DPs (acting as identifyees) we predict a different situation. Since it is not necessarily presupposed that the relevant discourse referent has already been introduced we expect that the identifier can contain new information and therefore a focussed constituent. This prediction is borne out. Lexical DPs allow focus within the identifier as well as a focus particle associated with the identifyee whereas pronominals do not:

(36) (a) *Peter hat [die [Frau]$_{[F]}$ gesehen, [die$_x$ gerne BIER trinkt]$_x$*
Peter has the woman seen, who likes BEER drinking

(b) *Peter hat [das [Argument]$_{[F]}$]$_x$ vorgebracht, [daß MARIA gerne Bier trinkt]$_x$*
Peter has the argument presented, that MARY likes beer drinking

(c) **Peter hat es geglaubt, daß MARIA Bier trinkt*
Peter has (it) believed, that MARY beer drinks

(d) *Peter hat geglaubt, daß MARIA Bier trinkt*
Peter has believed that Mary beer drinks

(37) (a) *Peter hat [sogar die [Frau]$_{[F]}$ gesehen, [die$_x$ gerne Bier trinkt]$_x$*
Peter has even the woman seen, who likes beer drinking

(b) *Peter hat [sogar das [Argument]$_{[F]}$]$_x$ vorgebracht, [daß Maria gerne Bier trinkt]$_x$*
Peter has even the argument presented, that Mary likes beer drinking

(c) **Peter hat es sogar geglaubt, daß Maria Bier trinkt*
Peter has it even believed that Mary beer drinks

(d) *Peter hat sogar geglaubt, daß Maria Bier trinkt*
Peter has even believed that Mary beer drinks

4.1. Definites re-introducing a discourse referent

Notice however that I have argued that N provided by the lexical DP can also denote the set of already established discourse referent (all satisfying the property denoted in N). Thus a definite DP can also reintroduce an already established discourse referent (similar to what we have found in case of Identification of pronominals). In section III.6. we have already seen another (though related) difference between lexical NPs and pronominals which has to do with the degree of descriptive content. There I discussed the phenomenon of picking out an already established discourse referent which is not accessible anymore. There we have already seen that reintroducing a discourse referent is possible with definite DPs. The examples we discussed involved cases where already established discourse referents are not immediately accessible anymore. It was argued that by means of Identification the relevant discourse referent is reintroduces. We were solely concerned with coordination.

In this section I will show that there are other means of making discourse referents inaccessible by means of creating sets. Still the same phenomenon can be observed: A definite DP can pick out one particular member of this set. Within DRT there is an another operation that yield groups (besides summation). This is the operation of *abstraction*. Abstraction within DRT is needed in order to explain the possibility to anaphorically refer to quantified NPs. Consider the following example (taken from Asher 1993):

(38) *Many students in my class went on the camping trip. They enjoyed it.*

The operation of abstraction allows to form a group of all the objects satisfying $\alpha \cup \beta$. α corresponds to the first argument of the quantifier and β to the second. The group is denoted by the term $\Sigma x \alpha \cup \beta$, where Σ takes the DRS $\alpha \cup \beta$ and a discourse referent declared x in $\alpha \cup \beta$ and returns the group of all those objects y that satisfy the property $\lambda x \alpha \cup \beta$. Such a group is the intended antecedent of the pronoun in (38).

For the purpose of the present analysis I assume an extensional view of the denotation of the quantified DP. Thus assume that the quantified NP introduces a set of individuals X = (x,y,z). The individuals within this set are not accessible to anaphoric resolution, neither by a pronoun nor by an definite description as exemplified below:

(39) *Viele Studenten meiner Klasse fuhren auf Camping Urlaub. #Ihm/#Dem Studenten hat es gefallen.*
'Many students of my class went on a camping trip. #He/#the student enjoyed it.'

However, if an identifier is added (which should suffice to unambiguosly determine the relevant individual to be denoted) then the discourse above renders wellformed even with a singular definite DP:

(40) *Viele Studenten meiner Klasse fuhren auf Camping Urlaub. Dem Studenten von dem ich es am wenigsten erwartete, hat es am besten gefallen.*
'Many students of my class went on a camping trip. The student of whom I expected it the least, has enjoyed it most.'

The examples above show again that a definite DP can re-introduce an discourse referent. In these cases a set is introduced by the quantified NP, all members of this set have the property in common which is denoted by N. Therefore a simple definite DP that involves the same N is not felicitous for picking out one particular member of this set. If however and identifier is added, then anaphoric linking to the previously mentioned DP is possible. It refers to a singleton subset of the set denoted by the quantified N.

5. Split Antecedents

Since I assume that the identifier is associated with an index we expect that it is possible that identifiers can create a set via summation. Therefore it should be possible that the relative pronoun can have more than one accessible discourse referent as its antecedent. This is a prediction borne out by the well-known phenomenon of split antecedence (first discussed in Perlmutter & Ross 1970). Consider their example and the representation I am assuming:

(41) *[A man]$_x$ entered the room and [a woman] went out [[who]$_z$ were quite similar]$_y$.*
DP1: x = a man
DP2: y = a woman
DP3: z = (x,y)

Summation by a relative pronoun seems to be more restricted than it is for other instances of anaphoric resolution that creates a set. The following sentence (also

involving split antecedents) is not wellformed:

(42) **A man met a woman who were quite similar*

Notice that the difference between (42) and (43) is that the former involves coordination whereas in the latter the two intended antecedents occur within one single clause. For the moment let us leave this observation. I will come back to this issue in section VI.2, where I will be concerned with the locality constraint imposed on Identification.

6. Deriving another presupposition

There is another property of identified DPs our analysis can derive. Consider the following observation found in Bach (1974) (who cites Kasher for originally observing this phenomenon):

> *"that a restrictive relative clause presupposes the existence of entities of which the description given in the relative clause is not true. Thus 'the man that I saw' presupposes at least one man that I didn't see".* (Bach, 1974: 272)

Thus a sentence like (43) gives rise to a presupposition[17] in addition to its assertion:

(43) *Peter traf [$_{NP}$den Mann]$_i$, der gerne Bier trinkt.*
Peter met the man, who likes beer drinking

Assertion:	Peter met NP_i P^{18}
Presupposition:	There is some NP_j (NP_j is not NP_i) such that ¬P is true of NP_j, i.e. $[NP_j \neg P]$

In addition to the assertion (that there is a man the DP denotes) the presupposition arises that there are other men who do not have the property denoted in the relative clause. This fact can be derived by our analysis. Consider again the representation we assume:

[17] Here we are talking about pragmatic presupposition in the broad sense. It concerns the appropriateness of a given sentence in a certain situation/discourse

[18] Here NP_i refers to the NP *the man* and P refers to the (identifying) property denoted by the relative clause.

(44) *[der [Mann]$_X$]$_e$ [[den$_e$ [ich gesehen habe]]$_x$*
[the [man]], [[who] [likes beer drinking]

NP: X = the set of individuals satisfying the property of being a man
DP: x = ?
$x \subseteq X$
CP: Y = set of individuals satisfying the property they like beer drinking
y = male, singular
$X \cap Y = (x,y)$
x = y

The crucial part of the representation of the relative clause is that x has to be a subset of X (i.e. the set of men). I have claimed that the property provided by the identifier must suffice in order to pick out the intended individual. This results in the interpretation that there are other individuals (within the set denoted by X) that satisfy the property of being a man. However if Identification should be felicitous the identifier must pick out the relevant individual unambiguosly. Thus the property denoted in the relative clause must be a property the other individuals contained in this set do not have. Therefore the presupposition arises that there are other men (sets of individuals denoted by N) that do not have the property denoted by the identifier.

Suppose that the relevant (presumably previously established set) contains just one single man. Then this presupposition does not arise. However, then the relative clause is not to be considered as an instance of Identification. In this case the property denoted by the relative clause does not provide a condition that conveys information to pick out the relevant discourse referent. In this case the relative clause adds further information about a discourse referent which is unambiguously determined. This is what we find in case of non-restrictive relative clauses. Non-restrictive relatives clauses are not licensed as an instance of Identification. I will not address the issue of non-restrictive modification. Suffice it to say that it is not unexpected that non-restrictive relative clauses impose different licensing requirements. For example they seem to be subject to a different locality constraint: it is a well-known fact that a non-restrictive relative clause (as opposed to a restrictive relative clause) cannot occur in extraposed position.

As we expect Bach's observation is not only true for relative clauses but also for other instances of Identification. The following sentence exemplifies this prediction. In case of a complement clause identifying a DP we arrive at a similar presupposition:

(45) *Peter hat gestern das Argument präsentiert, daß Maria gerne Bier trinkt.*
Peter has yesterday the argument presented, that Mary likes beer drinking

Assertion:	Peter presented NP_i C[19]
Presupposition:	There is some NP_j (NP_j is not NP_i) such that ¬C is true of NP_j, i.e. $[NP_j \neg C]$

One has to be careful with judgements here. As usual with DPs denoting abstract entities things are more delicate. The presupposition that arises is not necessarily that there are other existing arguments (that do not have the content C = *'that Mary likes beer drinking'*). Rather the relevant presupposition is that there are other *possible* arguments that do not have the content C denoted by the identifier. However I do not think that this is a problem for the analysis suggested here. Remember that I have argued that the X associated with pronominals denotes the set of *possible* antecedents. Moreover, we will see other examples where DPs denoting abstract entities behave differently than DPs denoting concrete entities. There the same phenomenon will arise, namely that abstract entities give rise to the presupposition that there are other *possible* entities (that satisfy the property denoted in N).

7. Contrastive Focus

The phenomenon we are dealing with involves new information in two ways. On the one hand we said that the identifyee (re-)introduces a new discourse referent. On the other hand the identifier must provide a sufficient conditions in order for the intended referent to be picked out (of a given set, which can either be previously established or itself being introduced by N). The notion of new information is related to focus. In this section I will be considered with the behavior of Identification in interaction with (contrastive) focus. In doing so we will also address the observation discussed in section II.2.4., namely that stress on the definite determiner obligatorily triggers Identification.

[19] Here NP_i refers to the NP *the argument* whereas C refers to an identifying property denoted by the complement clause. In this case the identifying property equals the *content* of the denoted *argument*.

7.1. Stress on the determiner

Rooth (1985) analyzes contrastive stress along the following lines. Consider the following sentences and the representation Rooth (1992) associates with them:

(46) (a) $[\![_S$ [Mary]$_F$ likes Sue$]\!]^f$ ={**like** (x,s) | x ∈ E}, where E is the domain of individuals.

(b) $[\![_S$ Mary likes [Sue]$_F]\!]^f$ ={**like** (m,y) | y ∈ E}

The sentences differ as to which NP is focussed (i.e. *Mary* or *Sue*). $[\![\alpha]\!]^f$ is the focus semantic value for the phrase α in contrast to the ordinary semantic value $[\alpha]^o$. The focus semantic value corresponds to some set of alternatives from which the semantic value is drawn, or a set of propositions which potentially contrasts with the ordinary semantic value. The ordinary semantic value must however always be an element of the focus semantic value. It is not such an easy task to derive this set of alternatives. Intuitively we can again assume that it has to be salient in the discourse or provided by the sentence itself. Thus consider:

(47) *Ich habe MARIA gesehen (und nicht Hans).*
'I have seen MARY (and not John).'

If it is not clear from the context, i.e. if there is no previously established set to which the element under consideration is contrasted to, then the set of alternatives must be provided by the sentence itself. Again we observe apparent optionality. *und nicht Hans* has to be present if the set of alternatives is not provided by the context. This phenomenon is reminiscent to what we have said about Identification in general. The parallelism between Identification and contrastive stress is that both show an option as to whether the relevant information occurs in the sentence or whether it is already clear form the context. In case of contrastive stress it is the set of alternatives which might either be 'salient' in the discourse or introduced in the sentence. In case of Identification the denotation of a given NP might either be unambiguously determined or the exact conditions for identifying the discourse referent are provided sentence internally. Now consider again an example where a DP contains a contrastively stressed definite determiner:

(48) *Peter hat DAS Auto gekauft *(das bei Egon in der Werkstatt stand)*
Peter has THE car bought *(which at Egon in the garage stood)

As indicated the sentence is just wellformed if the identifier is provided within the sentence. Why should this be so? Ideally this phenomenon should follow from the analysis of contrastive stress in interaction with the analysis for Identification established here. Consider once more the (abbreviated) representation for the relevant nominals:

(49) NP: X = Set of entities satisfying the property of being a car (= set of alternatives)
DP: x = ?
$x \subseteq X$
CP: x = the car that was in Egon's garage

The representation I assume provides a set of alternatives (indicated by $x \subseteq X$, which corresponds to Rooth's claim that the semantic value must be an element of the focus semantic value). Here the set of alternatives is established by the DP (by means of the NP it contains). What is contrastively stressed is not the property but the determiner. The definite determiner has a uniqueness presupposition. Assume therefore that contrastive stress on the determiner focusses on this very property, namely *uniqueness*.[20] Informally the interpretation is as follows: the denoted individual (or entity) is unique among the other individuals (or entities) denoted by N (i.e. it is unique within this set). Since the set X is established by N and x is itself a member of X, it is not necessary to provide the set of alternatives additionally. Rather it is necessary to provide the property which makes x unique among the set of alternatives (X). This distinguishing property is provided by the identifier.[21]

Notice that what we have said so far is nothing else than what we have said for Identification in general: the identifier provides enough descriptive content to pick out the intended discourse referent of a given set. The difference here is that by means of stressing the definite determiner the uniqueness of the referent is focussed. Thus syntactically there has to be an identifier, in other cases the obligatoriness of

[20] Notice that English does not allow for stress on the definite determiner in the cases we are considering here (cf. Brugger & Prinzhorn (in preparation)). However a similar phenomenon can still be observed. Consider the following example:
(i) *Peter has seen THE girl*
Stress on the definite determiner induces a particular interpretation, namely that the girl (seen by Peter) was the most outstanding girl one can imagine. Thus there IS a particular uniqueness interpretation induced (similar to the interpretation of a superlative).

[21] I do not have anything to say as to why it is not the case that by means of stressing the definite determiner uniqueness is contrasted to *non-uniqueness* as one might expect.

Identification cannot be observed syntactically. We will return to the difference between contrastive stress on the determiner as opposed to other instances of contrastive stress in the next section. Let me just show that the same presuppositions arise as the ones we already have detected for definite DPs in general. First we expect that the identifier can contain new information (i.e. that it can contain a focussed constituent):

(50) *Peter hat DAS Auto gekauft, das bei EGON in der Werkstatt stand.*
Peter has THE car bought which at EGON's in the garage stood

Secondly we predict that the contrastively focussed DP can be associated with a focus particle (*sogar, nur*):

(51) *Peter hat nur/sogar DAS Auto gekauft, das bei EGON in der Werkstatt stand.*
Peter has only/even THE car bought, which at EGON in the garage stood

Thirdly Bach's observation is also relevant here. The presupposition arises that there is at least one other car which does not have the property of standing in Egon's garage. This is again a correct prediction. The prediction can be tested by means of exhausting the context. I.e. if we have a context with only *one car* available then focus on the definite determiner is not felicitous[22]

(52) *Bei Egon steht derzeit nur ein Auto. Peter hat #DAS/das Autotes gekauft.*
At Egon's stands momentarily only one car. Peter has THE/the car/it bought

Finally we expect the same facts to be found with other instances of Identification. The example below shows the effect with a complement clauses, which suffices to provide the distinguishing property by means of denoting the content of the *argument*:

(53) *Peter hat (sogar/nur) DAS Argument präsentiert, daß Maria/MARIA gerne Bier trinkt.*
Peter has (even/only) THE argument presented, that Mary/MARY likes beer drinking.

Finally, let me address the question of how stressed definite determiners differ from unstressed ones. As has been observed in Wiltschko (1994) there is a particular

[22] Notice the same should be true for Identification of definite DPs without stress on the determiner. However, in this case the relative clause can be interpreted as a nonrestrictive relative clause. Therefore adding a relative clause can still be felicitous.

intonation pattern associated with these sentences, which is not necessarily so in case of unstressed determiners: the intonation rises with the focussed element. Then it stays high until the extraposed constituent and it falls again once another focussed element occurs within the identifier. It seems to be the case that the presence of another focussed element in the identifier is obligatory (a fact I have been abstracting away in the previous discussion). Moreover, in many cases the examples under consideration are only felicitous under the following circumstance. It seems to be necessary that there is a set of entities which have the property denoted by N in addition to some other common property. This property is expressed by the identifier, however there must be one particular difference distinguishing it from the other elements within this set. This difference is the one denoted by the focussed element within the identifier. In the example above this would correspond to a context where we talk about *arguments that somebody likes to drink beer*, and Peter presented the very argument, *that MARY likes to drink beer*.[23] Assume that stress on the determiner indicates that the relevant set has already been established in the discourse, i.e. it seems to be the topic of conversation. However, in case of unstressed determiners the set can either be already established or it is only introduced by the noun (contained in the DP). This amounts to saying that focussing uniqueness (and this is what stress on the definite determiner indicates) is only felicitous if there is a given set and one particular member of this set differs from the others in a very particular way. I do not have much more to say about this phenomenon. The exact reason for this peculiar behavior is an open problem for the present analysis.[24]

7.2. Stress on N

There is a difference between stress on the determiner and other instances of contrastive stress (i.e. the cases mainly discussed in Rooth). The relevant difference is as follows. As we have seen in the last section, contrastive stress on the determiner (i.e. stressing uniqueness) results in obligatory Identification. The identifier provides a property that suffices to unambiguously determine the relevant discourse referent. As we have already seen other instances of contrastive stress differ. Consider the following examples:

[23] This observation is due to Gerhard Brugger (p.c.)

[24] Moreover, notice that adding a restrictive adjective (which can in principle also provide a distinguishing property) does not license stress on the determiner:

(i) *Peter hat DAS große Auto gekauft *(das bei EGON in der Garage stand).*
Peter has THE big car bought *(that at EGON's in the garage stood).

(54) (a) *Peter hat MARIA gesehen (und nicht den Hans).*
Peter has MARY seen (and not John)
(b) *Peter hat eine ZIGARETTE geraucht (und keine Pfeife).*
Peter has a cigarette smoked and not a pipe

No matter whether the discourse referent is unambiguously determined (in (54a) a proper name is contrastively stressed) or not (in (54b) the N within an indefinite NP is stressed) Identification is not obligatory. In both cases however we find the necessity for a set of alternatives to be established. The interpretation that arises is that the denoted individual is contrasted to this set of alternatives, i.e. it is the only individual of the superset that has a particular property denoted within the clause (i.e. being seen (by Peter), or being smoked (by Peter)). In other words the property denoted by N is the distinguishing property of the individuals under consideration (the set of alternatives) that contrasts it to the other members of the relevant set. Let us assume that the set of alternatives corresponds to the set of possible individuals (relativized to the present discourse[25]) similar to what we find in case of pronouns. This means that the denotation of X (associated with N) does not correspond to the set of individuals satisfying the property denoted by N (i.e. the set of cigarettes). This assumption captures the relevant interpretation. Stress on N indicates that it is the stressed property itself (i.e. being a cigarette) that corresponds to the distinguishing property. In addition another member of this set can be explicitly added (*and not....*). (This individual is also interpreted as a member of the set of possible discourse referents). According to this assumption the difference between stress on N and stress on D respectively is clear. If D is stressed, N denotes the set of alternatives. The distinguishing property has to be provided by an identifier. If N is stressed, the denotation of N is the distinguishing property of the denoted individual. Therefore N cannot denote the set of alternatives. The set of alternatives has to be already established. If however, N denotes the distinguishing property, then a different behavior w.r.t. Identification is expected. In this case the identifier cannot be interpreted as the distinguishing property of the individual to be denoted as in the examples discussed in earlier sections. Consider now an example where a potential identifier (i.e. a relative clause) is added:

(55) *Peter hat das AUTO gekauft, das bei Egon in der Werkstatt stand.*
Peter has the CAR bought, which at Egon in the garage stood

[25] It can either be explicitly established or contextually salient.

Here the potential identifier is not interpreted as a distinguishing property of a particular car rather it is interpreted as a common property of the set of individuals the relevant discourse referent is contrasted to. According to this assumption we expect that in this case adding a further member of the set of alternatives is necessarily interpreted as having the property denoted by the relative clause as well. Indeed this is what we find in the example below:

(56) *Peter hat das AUTO gekauft, das bei Egon in der Werkstatt lag, und nicht das Motorrad.*
Peter has the CAR bough, which at Egon in the garage stood, and not the motor bike.

Here *the motor bike* is necessarily interpreted as having the property of *standing in Egon's garage*. I think the explanation for the difference between contrastive stress on N and on D respectively shows that the identifier can not just identify x (i.e. DP) but it can also identify X, i.e. it can provide a further condition on the set of individuals the actual referent is a subset of. Consider the (abbreviated) representation of a sentence like (55):

(57) NP: X = the set of entities satisfying the property of being a car.
DP: x = the car
CP: Y = set of entities satisfying the property of standing in Egon's garage

This representation does not give us the relevant interpretation. It does not contain any incomplete condition for the identifier to complete. Intuitively the relevant representation must include some X = ? indicating that N does not provide enough content. The identifier provides a property of the superset out of which the actual referent is picked. Remember that the property denoted by X serves as the distinguishing property, i.e. the property that contrasts the denoted individual to some other individuals. Although the set of alternatives can already be established it can still be the case that the relevant superset needs a further condition. Notice that (according to the present line of reasoning) it is also relevant to represent the presupposition that the denoted individual is a subset of the individuals satisfying the property denoted in the relative clause. We have to include a subset relation that captures this insight. Given these necessities I would like to propose the following representation for the example above:

(58) NP: X = the set of possible discourse referents
X = ?
CP: Y = the set of entities satisfying the property of standing in Egon's garage.
y = neuter, singular
$X \cap Y$
DP: x = ?
$x \subseteq (X \cap Y)$
Focus:
NP: X = the set of entities satisfying the property of being a car
x = y

This representation captures the relevant properties of contrastive stress on N we are interested in here. Let me briefly explain it. First of all it reflects a general assumption concerning the representation of focus. Focus can be analyzed in terms of (A'-) movement; the focussed constituent is moved and leaves a variable (among others, this view is taken in Chomsky 1976,1981, Huang 1982, Horvath 1986 and Rochemont 1986). Within the representation I assume this assumption is reanalyzed by claiming that the first instance of NP contains an incomplete condition ($X = ?$), which amounts to saying that N is somehow 'emptied'.[26] It is not before "Focus" is represented (i.e. the second instance of NP above) that this condition is completed (similar to binding the variable within the movement-approach). The focussed property is interpreted as the distinguishing property of the relevant discourse referent. If the denotation of X equates the distinguishing property, then it is clear there is no property left that can be taken as the common property of the superset. Secondly the representation in (58) is also reminiscent of the representation of pronominals as introduced in chapter III. In case of pronominals, X corresponds to the set of possible antecedents. X (since it is empty) can correspond to an arbitrary (salient) set of possible antecedents (like in case of pronominals) but it can be further identified. The identifier provides a further restriction on X. The denotation of the relative clause should suffice to pick out the relevant subset.[27] This means that we have to create the intersection of X and Y. The new set then is interpreted as the superset for the individual to be contrasted to (i.e. the set of alternatives). Given this interpretation we can conclude that a potential identifier can either identify DP (x) or NP (X). Notice that NP is under normal circumstances

[26] It might be that at LF N moves out of its position, i.e. there might be a 'Focus-position' within the DP. However I will not discuss such a possibility of Focus as an instance of LF (or LF')-movement.

[27] Notice that if X and Y would correspond to the same set, (i.e. if X = Y) then the identifier would be interpreted as a nonrestrictive relative clause.

not available stress on N seems to be a possibility to make X accessible for Identification.[28] However, if both possibilities are in general allowed the following question arises: Why is an identifier necessarily associated with N (X) if N is stressed? One might expect that in case of stress on N identification of DP is still a possibility. I do not have a real explanation for this property. However, I do not think that this causes a particular problem for the present analysis. That focus on N indicates that the identifier is necessarily associated with X is not really surprising. Notice that stress has this property not only in case of Identification. It is a well known fact that focus particles like for example *sogar* ('even') necessarily associate with focus. If no particular stress is added then it associates with its sister constituent. However, stress can render this (default) domain of association. If a sentence contains both, *sogar* and a stressed constituent, then *sogar* is necessarily associated with the stressed constituent, no matter in which position it occurs.

(59) *PETER hat die Maria sogar geküßt.*
Peter has the Mary even kissed.
'Even Peter has kissed Mary.'

In the example above *sogar* necessarily associates with *Peter* rather than with *geküßt*.

Another crucial property of Identification of X is captured by the representation above: The denotation of the relative clause is identified with the denotation of N (X = Y). The new information (focus) here is that it is the denoted individual (as opposed to somebody else) who Peter has seen. The superset is already established - it corresponds to X (the set of possible individuals). The identifier only restricts this superset. Since the superset is already established, then it must be the case that the various properties of its subsets are also in the background knowledge (i.e. they are already established/salient). Therefore we expect that the property denoted by the identifier is presupposed. As a consequence, it cannot contain new information, i.e. it cannot contain a focussed phrase:

(60) (a) **Peter hat die FRAU gesehen, die gerne BIER trinkt (und nicht den Mann*
*Peter has the WOMAN seen who likes BEER drinking (and not the man)

[28] In chapter VII. we will see that there is another possibility to doing so: I will argue that the (topicalized) d-pronoun in German Left Dislocation has the same effect: it makes N (X) available.

(b) **Peter hat das ARGUMENT präsentiert, daß MARIA gerne Bier trinkt (und nicht den Beweis).*
Peter has the ARGUMENT presented, that MARY likes beer drinking (and not the proof).

Notice that we have already found two different 'presupposition patterns'. The presupposition arising in case of stress on N patterns with the presupposition found in case of Identification of pronominals: the denotation of the identifier has to be presupposed. Stress on D behaves differently (as discussed in the previous section): the identifier must contain new information. As we have seen this difference can be captured by the present analysis. For the moment let me leave this analysis as we postulated it. We will return to this issue in chapter VI. and in the excursion on *einzig-NPs*. Let me briefly summarize the main findings. The most important point is that Identification is in principle available for x and X (i.e. for NP and DP). There are two important consequences. This suggests that there are in principle two possible adjunction sites for the identifier available: NP or DP.[29] We will see that this assumption makes predictions that are empirically supported. The second consequence is that an NP (i.e. a property) is associated with a discourse referent. I have to assume this because of the following reason. The main claim of the present analysis is that the sentential constituents under considerations are licensed as identifiers. Identification is a relation that holds at domain D. Since we found that X (i.e. NP) can be identified we can conclude from the previous discussion that X can serve as an entity for domain D. X corresponds to the denotation of NP, i.e. it denotes a set of individuals (if extensionally defined). The conclusion that X is an entity at domain D amounts to saying that properties (like N) are associated with a discourse referent. In the next section I will discuss some data that lead to the same conclusion. They will provide us with independent evidence for the validity of the claim that properties can be associated with discourse referents.

[29] Notice that this conclusion differs from standard analyses of relative and complement clauses. It has been a longstanding claim that a restrictive relative clause is within the scope of the determiner whereas nonrestrictives are not (cf. Jackendoff (1977) among others). Various different considerations lead to the present conclusion: there is a distinction between Identification of D and N respectively: 1) identifiers are necessarily in a position higher than the identifyee in order for the former to license the latter (cf. chapter I). 2) Facts concerning the locality constraint on extraposition will support this conclusion (cf. chapter V.). 3) Also within an analysis that takes the relative clause to be generated within the DP it has to be assumed that it can adjoin to DP. This is necessary in order to account for the fact that DP does not induce an island for 'extraposition' and to analyze 'DP-internal extraposition'. If however, DP-adjunction is an option for the DP-internal analysis of relative clauses, it is hard to see empirical differences between the two approaches in general. I will continue to assume that a relative clauses is basegenerated in a position adjoined to DP.

V. PREDICATIVE AND QUANTIFIED NPs

The conclusion we arrived at in the last section seems to contradict standard treatments of predicative NPs, which can be interpreted as properties. Under standard assumptions within DRT it is assumed that predicative NPs do not introduce discourse referents. On the other hand our analysis forces us to assume that this cannot be a valid treatment of properties in general. In the last section we have seen evidence that identifiers CAN be associated with properties (i.e. they can identify N). The same conclusion will equally apply to quantified NPs,[1] which will receive a similar analysis. Let me briefly outline the argumentation leading to the conclusion that properties are associated with discourse referents.

1. Properties as discourse referents

If the obligatoriness of *right*-adjunction is reduced to a precedence constraint, which itself is reduced to the Novelty Condition, it crucially depends on the assumption that the identifyee is associated with a discourse referent. Under standard assumptions this analysis creates a problem given the fact that it is possible to have a predicative as well as a quantified NP construed with a relative clause or a complement clause (i.e. with an identifier):

(1) (a) *Hans hält Maria für eine Frau, die gut kochen kann.*
John considers Mary a woman who well cook can
(b) *Hans kennt keine Frau, die gut kochen kann.*
John knows no woman, who well cook can
(c) *Hans kennt viele Frauen, die gut kochen können.*
John knows many women who well cook can

(2) (a) *Hans hält die Nudelsuppe für einen Beweis, daß Maria kochen kann*
John considers the noodle-soup for a proof, that Mary cook can.
(b) *Hans fand keinen Beweis, daß Maria kochen kann*
John found no proof that Mary cook can
(c) *Hans fand viele Beweise, daß Maria kochen kann*
John found many proofs that Mary cook can

[1] Remember that we have already argued that quantified NPs are associated with a discourse referent in section IV.4.1. There I discussed the operation of abstraction. In this section we will be concerned with some related phenomena.

The problems I have to address concerning those sentences w.r.t. the analysis I suggested are as follows

1) it is usually assumed that predicative NPs as well as quantified NPs do not introduce discourse referents
2) especially in (1b) there is no particular woman (in domain D) that the relative could identify. To the contrary it is asserted that there is no such discourse referent.

I will show in this section that these two apparent problems are related. The necessary conclusion we have to draw if the analysis suggested here is on the right track predicts some interesting results w.r.t. both predicative and quantified NPs construed with identifiers. Here is the conclusion I have to draw on basis of the examples in (1):

> In predicative as well as quantified NPs (i.e. non-denoting/non-referential nominal elements) the identifier necessarily identifies X (N), i.e. the property of the set denoted by N.

Let me first discuss the standard assumption within DRT, i.e. the reason for claiming that predicative NPs are not associated with discourse referents.

1.1. Predicative NPs in DRT

It is assumed for example in K&R that predicative NPs are not associated with a discourse referent. They cite Doron (1983) for the following evidence supporting this assumption. In Hebrew it is impossible to refer anaphorically to a noun phrase following the verb *be*. Consider the following quotations from K&R:

> *"... in sentence pairs of the form*
> *The N is (a) N'. He.../*She...*
> *where N is masculine and N' feminine, only the masculine pronoun can be used to pick up the individual mentioned in the first sentence. If the part after 'is' were to introduce its own discourse referent y it would no longer be explicable why the feminine pronoun, which fits the gender constraints on y, could not be used in this context."* (K&R, 1993: 268f)

> *"If, however, the part following 'is' functions as a predicate, no such discourse referent will get introduced."* (K&R, 1993: 269)

The same argumentation could be extended to quantified NPs. Consider the following examples:

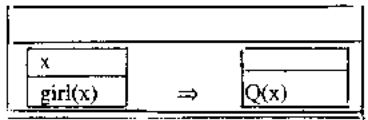
(3) (a) *#Jeder Mann will Maria küssen. Er kennt sie noch nicht einmal.*
'#Every man wants to kiss Mary. He doesn't even know her.
(b) *#Kein Mann will Maria küssen. Er kennt sie noch nicht einmal.*
'#No man wants to kiss Mary. He doesn't even know her.'

The sentences in (3) indicate that it is not possible to anaphorically relate a pronoun to a quantified NP. If the possibility for anaphoric reference is interpreted as indicating that no discourse referent is available then we could draw the conclusion that quantified NPs do not introduce discourse referents as well. Notice however, that in this case DRT assumes that there IS a discourse referent associated with the quantified NP. However, since it is in a subordinate position it is not immediately accessible for anaphoric linking by a pronoun. This line of reasoning is also taken in Asher (1993) for quantified NPs. Consider the following sentence:

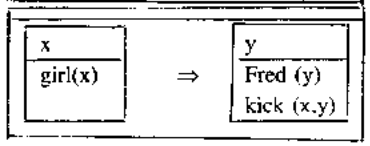
(4) *Every girl kicked Fred*

The partial DRS for the quantified NP *every girl* is given in (5). In combination with the translation of the VP this yields the DRS in (6):

(5) λQ



(6)



(Asher 1993: 75)

Asher assumes that:

> *"A discourse referent is also introduced in this case, but since it occurs in an embedded DRS [...] it cannot be referred to in subsequent discourse."*
> (Asher 1993: 75)

This means that we are back to the problem of accessibility, which I will not discuss any further arguing for a hierarchical organization of discourse representation (cf. section III.6.).

Notice that apparently the two cases of the impossibility to anaphorically link a pronominal to a given NP are treated differently within DRT. For predicative NPs the impossibility to anaphorically relate to a given NP is interpreted as to indicating the absence of a discourse referent. In case of quantified NPs the impossibility for anaphoric linking is interpreted as non-accessibility of the discourse referent. Both line of reasonings are relevant for the analysis of predicative NPs. First we can simply assume that the impossibility of anaphoric reference in the example discussed in K&R indicates that the discourse referent is NOT accessible. Second it seems to me that the test for the presence of a discourse referent on basis of anaphoric reference is valid (but only in one direction). If we find an instances of a pronoun that is anaphorically related to a predicative NP then we have evidence that predicative NPs are equally associated with discourse referents. The following sentence shows that anaphoric linking to predicative NPs is indeed possible:

(7) *Peter ist [ein Angsthase]$_i$. Hans ist [dasles]$_i$ auch.*
Peter is a coward. John is that/it too.

The pronoun in (4) is anaphorically related to the predicative NP *'ein Angsthase'*. I therefore conclude that a discourse referent is indeed available. What could this discourse referent be? It is obvious that it is not an 'existing individual in the world'. It is however a property which is ascribed to Peter in the first sentence. Thus what we can conclude is that:

(8) *Properties are associated with discourse referents.*

This means that considerations about anaphoric relations lead to the same conclusion we arrived at in the last section, a conclusion the analysis suggested here forces us to assume. Notice moreover that the line of reasoning K&R take for predicative NPs is not valid. If a pronoun X in a given context cannot be anaphorically related to an NP Y then two possibilities for interpreting this impossibility arise: 1) Y is not associated with a discourse referent or 2) Y is associated with a discourse referent y, but y is not accessible for anaphoric resolution. This amounts to saying that the possibility for anaphoric reference implies that there is a discourse referent available. However, the opposite implication is not true, i.e. the impossibility for anaphoric reference does not imply that there is no discourse referent. Notice that for predicative NPs (but apparently not for quantified NPs) K&R assume the implication to be true in both directions, i.e. they take the first line of reasoning above without taking the second one into consideration. Here we will be considered with this second possibility.

Let us draw attention to the nature of the discourse referent introduced by predicative (and quantified) NPs. Remember that we made use of an extensional definition for properties. Therefore I will assume for predicative NPs that the discourse referent they establish corresponds to the set of individuals satisfying the property denoted by N. This is similar to V&Z's type-interpretation as discussed in section II.2.1.1. They give the following evidence for the fact that the type interpretation is available (i.e. that it is associated with a discourse referent). As expected it can induce a plural interpretation (remember the difference between sets of individuals and sets of facts/events, the former being in general referred to by plural pronominals, cf. section III.6.5.1.). Consider the following examples taken from V&Z:

(9) *Parce que leurs griffes peuvent devenir acérées avec l'age, Marie a décidé d'acheter des chaussettes à ce chaton*
Because their nails can become very sharp with age, Marie decided to buy socks for this kitten

(10) (a) *Because they are usually well illustrated, every child bought a book on animals.*
(b) *Because they are usually well illustrated, John decided to buy the book on animals on the shelf.*

Notice that these data provide an important piece of evidence for the present claim. As we have seen in section III.5. backwards pronominalization is only licensed if the content of the pronoun is presupposed. It is an indication for an already established discourse referent. The above examples therefore indicate that properties are available as discourse referents. Thus it is not an 'individual-denoting' DP *the book* rather it is the property (type) denoting NP *book* which is denoted by the pronominal. Since it refers to a whole set (the set of individuals satisfying the property of *being a book (on animals)*) the pronoun referring to it occurs in plural.

Although I do not want to go into a fullfledged analysis, since this would go beyond the scope of this dissertation some interesting issues arise. Thus the following remarks are to be seen as positing some questions to be addressed in future research.

2. 'Hidden partitivity' in quantified NPs

These remarks imply that quantified NPs can be accessible for anaphoric resolution in a similar way, i.e. the NP of the quantified NP is expected to be available for anaphoric resolution. Since we claim that it is the property (extensionally defined as

a set of individuals) which introduces the discourse referent we predict that the adequate pronominal can occur in plural. Consider the following sentence:

(11) *Keine Frau hat Hans geküßt. Sie fanden ihn (alle) zu abstoßend.*
No woman has kissed John. They (all) considered him too disgusting.

Anaphoric relation is possible with a plural pronoun. The intended interpretation is partitive. W.r.t. domain D the sentence is interpreted such that *none of the women relevant for the present discourse kissed John*. Thus there is a *set of women* and it is asserted that *none of them kissed Mary*. The pronoun in the subsequent clause refers to this *set of women*. Notice that this line of reasoning is reminiscent of Enç's (1991) claim that quantifiers are inherently specific in the following sense. It is never the case that they quantify over all possible individuals in the world but just over contextually given sets. Enç (1991) claims that the following two sentences are equivalent:

(12) *Sally danced with every man.*
(13) *Sally dances with one of the men.*

She then concludes:

> *"that contextually relevant means 'already in the domain of discourse', since the contextually relevant individuals are those that have been previously established in the discourse. If universal quantification is over contextually relevant sets of individuals, it follows that NPs that quantify universally are specific. This account also ensures that universally quantifying NPs presuppose existence.'* (Enç, 1991: 11)

I think that these remarks support the claim that domain D should be assumed as an independent level of representation where discourse related facts are encoded. The representation of a sentence like (12) assumed to hold at LF is as follows:

(14) For every x, if x is a man, then Sally dances with x.

The covered partitivity cannot be read off this representation. The representation of nominals at domain D as assumed here is as follows:

(15) For every x, $x \subseteq X$, (X is the set of man), Sally dances with x.

It is obvious that the denotation of the NP induces the restriction, i.e. it corresponds to the restrictive clause in standard logical treatments. Since at domain D is assumed to represent the denotations relevant for discourse the intended (hidden partitive)

interpretation follows. Furthermore as opposed to Enç we do not have to say that the inherent partitivity implies previously established individuals. It can also be the case the X is introduced in a sentence like (12). The inherent partitivity arises by means of $x \subseteq X$, where X is the set of men, relativized to the present context. These remarks suggest that the notion of restrictive clause vs. nuclear scope (as assumed in standard treatments of quantification) is not just relevant for clauses (IPs) but also DP-internally. The representation assumed here defines a restrictive clause for every DP (i.e. NP always induces a restriction on the denotation of DP). In the excursion on *einzig*-NPs we will see that this is a desired result. There I will discuss some phenomena reminiscent of the tripartite structure of clauses that show that this notion is also relevant DP-internally.

It should also be clear now that the problems addressed at the beginning of this chapter are related. Consider again the relevant sentences repeated here for convenience:

(16) (a) *Hans hält Maria für eine Frau, die gut kochen kann.*
John considers Mary a woman who well cook can
(b) *Hans kennt keine Frau, die gut kochen kann.*
John knows no woman, who well cook can
(c) *Hans kennt viele Frauen, die gut kochen können.*
John knows many women who well cook can

I will assume that the relative clause identifies N, which is associated with a discourse referent (rather than identifying D). So I conclude that there is no problem for the present analysis induced by the fact that (16b) it is asserted that there is no actual discourse referent.

3. Accessibility of X

Having established the claim that NPs (i.e. expressions which denote properties) can establish a discourse referent we have to address the question as to when these NPs are accessible as discourse referents. Two possibilities arise. Either one claims that a predicative NP is categorically just an NP and thus it just establishes the index X, i.e. no x is introduced. Thus if K&R talk about 'no discourse referent' we are only talking about 'no x'. However, X is (always) introduced and it can act as a discourse referent. Alternatively we could say that a predicative NP is categorically a DP, with an 'expletive' determiner (in the sense of V&Z) or with a determiner that does not bind

the variable provided by N (in the sense of Higginbotham (1985) or Stowell (1987) respectively). (Cf. section II.2.) for a discussion of the two possibilities).

One hypothesis concerning the accessibility for anaphoric resolution (which is syntactically motivated) would be to claim that once an NP is dominated by a DP (and therefore (somehow) closed (the argument of N being bound by the determiner) then the discourse referent (X) is not accessible anymore. As we have already seen members within a 'closed set' are not (easily) available for anaphoric linking. But this situation can be changed in case the anaphorically linked element has enough descriptive content to pick one member out of this set.

Let us return to the example in (7). K&R's example from Hebrew apparently shows that no anaphoric relation is possible in case of predicative NPs. The German example however shows the opposite. The difference between K&R's example and the one in (7) has to do with the nature of the pronoun. In the Hebrew example we find a feminine NP and a feminine pronoun. The German example involves a masculine N and a *neuter* pronoun. It seems to be a quite plausible assumption that properties in general do not have gender features. Given the possibility that we can interpret neuter to indicate the lack of gender features, then it is not surprising that a female pronoun cannot anaphorically link to a predicative NP (i.e. a property). The neuter pronoun has to be used instead. Similar effects will be seen in chapter VII., where I will discuss Left Dislocation. First it will be seen that pronominals are able to make X accessible (a similar mechanism as stress on N) and second we will see some other instances of properties anaphorically related to a neuter pronominal.

However this discussion imposes a problem. In the sentences in (16) a predicative or a quantified NP is identified. However, the relative pronoun is singular and feminine and neither in plural nor in neuter, i.e. it agrees in gender with the head-NP. This fact seems to contradict the above conclusion, thus it poses a problem for the claim that properties are to be defined extensionally as sets and thus anaphoric relations are just possible if the pronominal is a neuter or plural one (like in the other cases we discussed). Why should this be so? I think that the answer to this problem lies in the difference between anaphoric relations by means of a pronoun and coreference by means of a relative pronoun. The latter are to be found within a more local domain, i.e. their accessibility seems to be less restricted. Remember that I am assuming that in case of Identification of NP the identifier is adjoined to this NP. In this position the phi-features of the NP under consideration are still accessible. Moreover the identifier is only licensed by being mapped to the identifyee. It is not inherently licensed. However, other instances of anaphorically linked pronouns are inherently licensed in

the position they occur in. Apparently this makes a difference for the nature of the pronoun. There is one more interesting issue to address, which I think is related but I cannot offer a solution. Again it concerns the relative pronoun. In general properties allow for non-restrictive relatives. But these relative pronouns occur as neuter wh-pronouns:

(17) *Marias Hund ist bissig, was sich für einen Hund eigentlich nicht gehört*
'Mary's dog is snappish, which is not appropriate for a dog.'

The relative clause is construed with the property (VP) 'being snappish'. The same phenomenon can be observed with NPs:

(18) *Hans sucht eine Sekräterin, was Maria zufällig ist.*
John is looking for a secretary, which Mary happens to be.

Here the pronoun is correlated with the NP, but it does not agree in gender, it is neuter. Now consider some other instances, where a neuter wh-pronoun occurs in the relative clause:

(19) (a) *Wir haben nichts, was/*das wir essen könnten*
We have nothing, what/d-pron we eat could
(b) *Das ist das Beste was/*das wir finden konnten*
This is the best what/d-pron we could find.

There are several nominals that (when associated with a relative clause) only allow for the wh-pronoun. One instance is for example the superlative. However, if the superlative receives a [+human] denotation this possibility disappears:

(20) *Das ist der Beste, den/*was/*wen wir finden konnten.*
This is the best, who/what/d-pron we find could

So it seems that within relative clauses the phi-features (as well as the feature [+human] of the head NP is still available. At the moment I do not have anything of interest to say about the above paradigm. It indicates at least that there is something more complex going on as would be implied by the previous discussion.

The main purpose of this chapter was to show that the possibility to identify predicative and quantified NPs does not cause a problem for the present analysis. We have seen that they can be associated with discourse referents. Some related issues will

be discussed in more detail in the excursion on *einzig*-NPs. Before doing so I will discuss some of the syntactic properties of Identification.

VI. SYNTACTIC PREDICTIONS

1. Predictions of the Precedence Constraint on Identification (PCI)

In this section I will discuss the syntactic consequences of the Precedence Constraint. I will first show the empirical facts it explains. Since one of the aims within syntactic theory during the last decades was and is to get rid of the notion of *precedence* I will also address this issue. I will discuss some other proposals (not involving reference to precedence) to derive the empirical facts I am discussing and show some problematic aspects inherent to these solutions.

1.1. The identifier cannot be left-adjoined to the identifyee

The first fact that can be derived is that an identifier can only be adjoined to the right and not to the left of the identifyee as shown in (1). Otherwise, the identifier would precede the identifyee:

(1) *[YP$_i$ [XP$_i$]] (where XP = identifyee; YP = identifier)

This holds for all the core cases we have been discussing so far, i.e. identification of pronominals[1] and lexical NPs (by a complement or a relative clause) respectively. Consider:

(2) (a) **Peter hat, [daß Maria Bier trinkt]$_i$, [daran]$_i$ geglaubt.*
Peter has that [Mary beer drinking]$_i$ thereon$_i$ believed
(b) **Peter hat [die$_i$ Bier trinkt], [die Frau]$_i$ getroffen.*
Peter has [who$_i$ beer drinks] the woman$_i$ met
(c) **Peter hat [daß Maria Bier trinkt]$_i$ [das Argument]$_i$ vorgebracht.*
Peter has [that Mary beer drinks]$_i$ [the argument]$_i$ presented

[1] Notice that the paradigm in (1) does not include an example involving the pronominal *es*. The prediction that the identifier cannot be left-adjoined is vacuous here. It is in general impossible to adjoin to a pronominal element. Thus it is also very marginal to right-adjoin to a pronominal. However there is still a contrast, left-adjunction being degraded:
(i) *?*Peter hat es daß Maria gerne Bier trinkt geglaubt.*
Peter has it that Mary likes beer drinking believed.
(ii) **Peter hat daß Maria gerne Bier es trinkt geglaubt.*
Peter has that Mary likes beer drinking it believed

Notice that so far I have not been explicit about the *base-position* of the identifier. I am assuming that the position right-adjoined to the identifyee is the base-position (in chapter VII. I will give indirect evidence for this claim). This is not the only possibility: there are analyses to be found in the literature that propose that the sentential elements classified as identifiers are base-generated within the complement position of their head N. However, the paradigm above is also a problem for those analyses. Under a DP (or NP-internal) approach for the position of these clauses the position of the identifier occurring to the right of the head would differ from the position to the left in that the latter would be a derived position (since there is no complement position available to the left of N). But these analyses have to admit for the possibility to right-adjoin to the whole DP because of two reasons.[2] The first reason is a theory-internal one: in order for the identifier to move out of the DP (to be extraposed to a position following the verb in final position) it would have to adjoin to DP in order to circumvent a subjacency violation. The other reason is an empirical one. It can be shown that *constituent-internal* extraposition is possible. Thus (as we will discuss in section VI.2.3.) it is possible to have an identifier construed with an DP which is embedded where another DP (presumably in complement position) linearly intervenes:

(3) (a) *Peter hat die Tochter$_i$ des Mannes, [die$_i$ Bier trinkt]$_i$ gesehen.*
Peter has the daughter the$_{GEN}$ man [who beer drinks] seen
(b) *Peter hat das Argument$_i$ des Mannes, [daß Maria Bier trinkt]$_i$ präsentiert.*
Peter has the argument the$_{GEN}$ man [that Mary beer drinks] presented

Here PCI equally rules out left-adjunction:

(4) (a) **Peter hat [die$_i$ Bier trinkt]$_i$ die Tochter$_i$ des Mannes, gesehen.*
Peter has who beer drinks the daughter the$_{GEN}$ man, seen.
(b) **Peter hat [daß Maria Bier trinkt]$_i$ das Argument$_i$ des Mannes präsentiert.*
Peter has [that Mary beer drinks] the argument the$_{GEN}$ man presented.

I do not assume that extraposition is derived via movement. (For my proposal it does not really matter whether the identifier can occur in complement position as well. In either case 'extraposition' would not leave a trace). If one wants to explain the

[2] cf. Müller (1994), who assumes the DP-internal analysis for the constituents under consideration. There the two reasons for allowing right-adjunction (as opposed to left-adjunction) are explicitly discussed. I will discuss his approach in Appendix I.

paradigm in (2) on basis of the assumption that there is simply no position to the left available then one has to explain why extraposition is allowed (even constituent internally to the right but not to the left). The PCI derives these facts straightforwardly. The only possibility to rule out the sentences in (2) and (4) without the PCI would be to stipulate a prohibition against left-adjunction (cf. Müller 1994). But leftness or rightness respectively should not differ hierarchically. Moreover talking about rightness implies talking about precedence. I take this paradigm to be direct support for the analysis presented here, i.e. that the ungrammaticality is derived by a prohibition for the identifier to linearly precede the identifyee (which reduces to a violation of the Novelty Condition).

1.2. The identifier cannot move to a position preceding the identifyee

The second prediction is that the identifier can never be moved to a position which precedes the identifyee:

(5) $*YP_i ... XP_i .... t_i$ (where XP = identifyee: YP = identifier)

It is important to notice that PCI applies to the linear output, i.e. to the S-structure position. Therefore reconstruction can not save PCI-violations. There are two empirical paradigms to be derived. The first one involves topicalization, i.e. movement to SpecCP the second involves movement to a position that would in principle be available for adjunction. Again it is only possible to right-adjoin the identifier. Consider the first paradigm. Movement to SpecCP is possible for any constituent in general (given that German is a V2-language). The examples below show that complement CPs can only be moved into that position if there is no correlate in the matrix clause:

(6) (a) *[Daß Maria Bier trinkt]$_i$ hat Peter (*es$_i$) geglaubt.*[3]
[that Mary beer drinks]$_i$ has Peter (*it$_i$) believed
(b) *[Daß Maria Bier trinkt]$_i$ hat Peter (*daran$_i$) geglaubt.*
[that Mary beer drinks]$_i$ has Peter (*thereon$_i$) believed
(c) **[Die$_i$ Bier trinkt]$_i$ hat Peter [die Frau]$_i$ getroffen.*
[who$_i$ beer drinks] has Peter [the woman]$_i$ met

[3] As we will see later on, pronominals that can be expletive at domain D (in the sense discussed in section III.3.) are not subject to PCI as expected: here the precedence constraint cannot apply.

(d) **[Daß Maria Bier trinkt]$_i$ hat Peter [das Argument]$_i$ präsentiert.*
[that Mary beer drinks]$_i$ has Peter [the argument]$_i$ presented

The second paradigm is the following one. In principle there is no obvious reason why an identifier which can be adjoined to the right cannot be adjoined to the left (where it would still obtain the Locality Constraint). In other words, if it were not for the PCI we would expect the following sentences to be wellformed:[4]

(7) (a) **Peter hat [daß Maria gerne Bier trinkt]$_i$ gestern es$_i$ geglaubt.*
Peter has [that Mary likes beer drinking] yesterday it believed.
(b) **Peter hat [die$_i$ gerne Bier trinkt]$_i$ gestern die Frau$_i$ gesehen*
Peter has [who likes beer drinking] yesterday the woman seen
(c) **Peter hat [daß Maria gerne Bier trinkt]$_i$ gestern das Argument$_i$ präsentiert*
Peter has [that Mary likes beer drinking] yesterday the argument presented.

For the sake of the argument let me assume (simplifying facts) that the pronominal is in the complement position of VP. The adverbial *gestern* could be in SpecVP or in a position adjoined to VP. Then the identifier would still be adjoined to the first maximal projection dominating (and thus i-govern) the identifyee. Notice that the intervention of *gestern* should not make any difference for the relation of Identification. If an adverbial intervenes to the right no violation of the Locality Constraint on Identification (LCI) arises as can be observed in the following examples:

(8) (a) *Peter hat es$_i$ geglaubt, bis gestern, [daß Maria gerne Bier trinkt]$_i$.*
Peter has it believed, until yesterday [that Mary likes beer drinking]
(b) *Peter hat die Frau$_i$ gesehen, vor nicht allzu langer Zeit [die$_i$ gerne Bier trinkt]$_i$*
Peter has the woman seen, before not alltoo long time who likes beer drinking
(c) *Peter hat das Argument$_i$ präsentiert, vor nicht allzu langer Zeit [daß Maria gerne Bier trinkt]$_i$*
Peter has the argument presented, before not alltoo long time that Mary likes beer drinking.

[4] Notice that the sentences in (2) and (4) could also be assigned a phrase-marker where the relative clause is left-adjoined to a projection of the main projection line. There is no basis on which we can decide between the two options - both are expected to be ruled out by the PCI anyway.

The sentences in (7) cannot be excluded as an instance of a violation of the LCI. The very same configuration is allowed if the identifier occurs in right-adjoined position. Again hierarchically the two sentences in (7) and (8) do not differ. There is no obvious reason why (7) is out while (8) is in. PCI can explain these facts.

1.3. A constituent containing the identifier cannot move to a position preceding the identifyee

The next prediction of the PCI is that no element containing the identifier can be moved to a position preceding the identifyee:

(9) *[$_{ZP}$.....YP$_i$]$_j$.....XP$_i$.....t$_j$
 |_____| (where XP = identifyee; YP = identifier)

The relevant examples involve VP-topicalization - a possible option in German in general. If the extraposed constituent is adjoined to VP then we expect that it is possible to move this extraposed constituent along with the VP. This is possible as can be observed in the following example:

(10) (a) *[daran$_i$ geglaubt [daß Eva Bier trinkt]]$_j$ hat Tim nicht t$_j$*
[thereon$_i$ believed [that Eve beer drinks]$_i$] has Tim not.
(b) *[es$_i$ geglaubt [daß Eva Bier trinkt]]$_j$ hat Tim nicht t$_j$*
[it$_i$ believed [that Eve beer drinks]$_i$] has Tim not.
(c) *[das Argument$_i$ geglaubt [daß Eva Bier trinkt]]$_j$ hat Tim nicht t$_j$*
[the argument$_i$ believed [that Eve beer drinks]$_i$] has Tim not.

Topicalizing the identifier along with the VP is only possible if the identifyee is also contained within the VP. As soon as the identifyee is left behind in the matrix clause the identifier cannot move along with the VP. In this case it would occurs in a position preceding the identifyee, violating the PCI. This is a correct prediction:

(11) (a) *[geglaubt [daß Eva Bier trinkt]]$_j$ hat Tim (*daran$_i$/*es$_i$) nicht t$_j$*
[believed [that Eve beer drinks]$_i$ has Tim (*thereon$_i$/*it$_i$) not.
(b) **[getroffen [die$_i$ Bier trinkt]]$_j$ hat Peter gestern [die Frau]$_i$ auf der Party t$_j$*
[met [who$_i$ beer drinks]]$_j$ has Peter yesterday [the woman]$_i$ at the party
(c) **[vorgebracht [daß Maria Bier trinkt]$_i$]$_j$ hat Peter [das Argument]$_i$ nicht t$_j$*
[presented [that Mary beer drinks]$_i$] has Peter [the argument]$_i$ not

Again the relevant generalization is as follows: as soon as the identifier occurs in a position preceding the identifyee the sentence is ungrammatical. Notice that this paradigm is also excluded by LCI. In (11) the adjunction site of the identifier (VP) would be lower than the position of the identifyee. LCI however predicts that the adjunction site of the identifier is higher than the position of the correlate (contra standard treatments of extraposition, cf. Appendix I.).

The paradigm above is a problem for most analyses of extraposition (e.g. C&R, 1990; Müller 1994) as well as for the rightbranching analysis (as I will discuss in Appendix I and II[5]). However it follows straightforwardly from the proposal suggested here.

1.4. No movement out of the identifier

The paradigms discussed so far lead to the following generalization: some instances of (legitimate) movement are blocked in case the output results in a configuration where a potential identifier precedes its (coreferent) identifyee. If no such overt coindexed element is present all the sentences are well-formed. Under the present account this generalization is derived by PCI. It is assumed that the presence of the identifyee in a position following the identifier is responsible for the ungrammaticality. The same generalization holds for another paradigm. This paradigm concerns extraction out of an identifier. Extraction out of an extraposed (complement) clause is generally possible as long as there is no correlate in the matrix:

[5] The problems are roughly as follows: Approaches which rely on some version of the complement principle have to face the problem as to why other complements can be moved quite easily out of their base-position (to a position where no government relation holds between the complement and the complement-taker). Haider (1993) - an advocate of the right-branching analysis claims that the ungrammaticality of the paradigm in (11) derives from a c-command requirement: i.e. the 'identifyee' must c-command the identifier. However the wellformedness of the following examples contradicts this assumption:
(i) *DIE Frau gesehen hat Hans, die gerne Bier trinkt.*
THE woman seen has John, who likes beer drinking
Here the identifyee is contained in the topicalized VP and the identifier is left behind in 'extraposed' position. It does not c-command the identifier, still the sentence is well-formed.

(12) *ZP_i.....XP_j.....$[_{YP\ i}$.....t_j.....$]_i$
|_____...._____| (where XP = identifyee; YP = identifier)

(13) (a) *Wer$_j$ hast du (*daran$_j$/*es$_j$) geglaubt [daß t_j Bier trinkt]$_j$?*
Who$_j$ have you (*thereon$_j$/*it$_j$) believed [that t_j beer drinks]$_j$
(b) **Was$_j$ hast du [die Frau]$_i$ gesehen [die$_i$ gerne t_j trinkt]$_i$.*
Who$_j$ have you [the woman]$_i$ seen [who$_i$ likes t_j drinking]$_i$
(c) **Was$_j$ hast du [das Argument]$_i$ geglaubt, [daß Maria gerne t_j trinkt]$_i$*
What$_j$ have you [the argument]$_i$ believed [that Mary likes t_j drinking]$_i$

The paradigm in (13) follows if we assume the following definition of precedence:

(14) XP precedes YP iff it linearly precedes every member of YP

The impossibility to extract is then derived in the following way: extraction creates a chain. The head of the chain is still a member of the extraposed element (YP). Extraction out of the identifier to a position preceding the identifyee results in a configuration where the identifyee does not precede every member of the identifier.[7]

[6] The following English counterexample was pointed out to me by Edwin Williams (p.c.):
(i) *Who did you see more often than Peter saw t.*
As I will argue in chapter VIII, comparatives are also instances of Identification. In sentence (i) wh-movement applies in an 'across-the-board'-manner resulting in a configuration where the wh-word ends up in a position preceding the identifyee (*more often*). I do not have anything to say about the well-formedness of the above example. It has to be noticed however that w.r.t. Identification English seems to work quite different to German, a matter which I will not address in this thesis. This comparative view on Identification is left open as a matter of future research. As far as the above counter-example is concerned it is relevant for the present discussion that German does not allow for this configuration, supporting the present analysis:
(ii) **Wen hast du öfter t gesehen, als Peter sah t*
The only well-formed German counterpart is the following example:
(iii) *Wen hast du öfter gesehen als Peter.*
In this example no chain is involved that would violate the precedence constraint.

[7] Notice that one might claim that there is another (non-structural) possibility to explain this fact. Remember that we said that the occurrence of a correlate induces a presupposed interpretation for the identifier. If the content is however presupposed then it might be argued that this is the reason for the impossibility to wh-move out of the identifier. But this kind of solution is not really an option because of two reasons. First it would contradict the autonomy of syntax hypothesis (which I am still maintaining although I make reference to discourse phenomena). We should better stick to a structural reason and the PCI is still a structural explanation (in the sense that it is not functional). Second if this line of reasoning would be on the right track it would be expected that at least topicalization out of the identifier should be possible, since there is no reason which would forbid topicalization of a constituent which is contained in a clause that is presupposed. However this prediction is not borne out:
(i) *Bier habe ich nicht (*daran) geglaubt daß Hans t trinkt*

Notice that I do not have any further motivation for this assumption. We would have independent support for the relevance of this definition of Precedence if we would find a configuration, where extraction out of an identifier is possible as soon as the landing site is in a position following the identifyee rather than preceding it. Long distance-scrambling would be the empirical domain to look at. Scrambling across a sentence boundary is very marginal in German in any case. However the following contrast seems to contradict the prediction of the precedence constraint:

(15) (a) **?Ich habe einen Pudding gesagt, daß ich esse.*
I have a jelly said that I eat.
(b) **Ich habe es einen Pudding gesagt, daß ich esse.*
I have it a jelly said that I eat.

Even in case the moved element occurs in position preceding the identifyee the presence of *es* results in a degraded output. Notice however that there is an additional problem with the sentence in (15b) that is captured by the present analysis. The LCI forces the identifier to adjoin higher as the identifyee. Therefore the identifier must be adjoined in a position dominating *es* and by transitivity it would also dominate the moved NP *einen Pudding*. Therefore the scrambled element would have to 'lower' in order to occupy a position following the identifyee. Lowering in general is excluded, since it results in a configuration where the antecedent does not c-command its trace anymore. So the sentence in (15b) is independently ruled out. It seems to be the case that the relevant example cannot be construed, since it would always result in an ECP-violation. Notice that the paradigm in (13) is a well known phenomenon which received several different analyses, which I will briefly discuss in the next section. This will then lead us to some general remarks concerning the use of a precedence constraint as opposed to analyses relying on hierarchical explanations.

1.4.1. Bennis (1986) and Cardinaletti (1990)

There are two main accounts for the contrast in (13). One is found in Bennis (1986) as well as in Cardinaletti (1990). Their claim is as follows: if the pronominal is present, then the extraposed clause has the status of an adjunct. Therefore it creates an island for extraction. If however the pronominal is not present, then the extraposed

beer have I not (thereon) believed that H. i drinks
This example shows that there must be a principled reason to exclude A'-movement out of the identifier which cannot be explained functionally.

clause is argued to be the argument of the verb. It does not induce an island. It is claimed in Cardinaletti (1990) that in this case the German VP 'turns into' a head-initial one. This is supposed to be the reason for the clause to be l-marked and therefore it does not count as a barrier for extraction. Notice that l-marking is the crucial factor to derive the argument- vs. adjunct status of the extraposed clause respectively. It can however be shown that there is a major empirical problem for such an approach. Extraction out of the extraposed clause (without a correlate) is even possible if there is another extraposed clause intervening between the verb and the 'argument' clause:

(16) (a) *weil Peter einem Mann$_i$ gesagt hat [den$_i$ er kannte][daß er Maria geküßt hat].*
since Peter a man$_i$ told has [who$_i$ he knew][that he Mary kissed has]
'...because Peter has told a man whom he knew, that he has kissed Mary'

(b) *Wen$_j$ hat Peter einem Mann$_i$ gesagt [den$_i$ er kannte][daß er t$_j$ geküßt hat]*
Who$_j$ has Peter a man$_i$ said [whom$_i$ he knew][that he t$_j$ kissed has]
'Whom has Peter told a man, whom he knows, that he kissed has'

(c) *Die Maria$_j$ hat Peter einem Mann$_i$ gesagt [den$_i$ er kannte] [daß er t$_j$ geküßt hat]*
The Mary$_j$ has Peter a man$_i$ said [who$_i$ he knew][that he t$_j$ kissed has]
'It was Mary that Peter told a man he knew, that he has kissed'

In these cases the 'complement CP' cannot simply be generated as a right-hand sister of V. However, in such cases of multiple extraposition extraction is still possible. I take these examples to be the major empirical evidence against an approach à la Cardinaletti (1990) or Bennis (1986). One way to solve this problem is related to the other account for deriving the contrast in (13) which I will discuss now.

1.4.2. Extraposition as a PF-phenomenon?

There are several different theoretical problems that lead to the conclusion that extraposition is best analyzed as a PF-phenomenon. One of these problems has already been mentioned in chapter II.1. Traditionally, apparently optional 'rules' are treated as PF-rules (or stylistic rules). This is also implicit in Chomsky's (1992) assumption that such rules are considered to be post-SPELL-OUT-processes mapping to PF. The second problem is more related to the paradigm in (13), it concerns the extractability vs. non-extractability out of extraposed clauses. Assuming extraposition to be a PF-

phenomenon seems to provide a solution. It can be assumed that wh-movement precedes 'extraposition' (since it takes place before S-structure - or SPELL OUT respectively). At the point of derivation when wh-extraction takes place the clause is still in its *base-position*. If the extraposed clause is an argument of the verb it is still in a position where extraction is possible (the clause being governed by the verb and therefore transparent for extraction). On the other hand the non-extractability out of extraposed clauses which are correlated with a nominal element in the matrix can be reduced to an instance of the Complex-NP-Constraint under the assumption that these clauses are really complements of N. I have already discussed problems concerning the latter assumption. There are also several arguments that extraposition is NOT a PF-phenomenon which I will briefly summarize here.

The first problem for the claim that extraposition is a PF-rule is specifically related to the paradigm in (13). It can be shown that assuming extraposition as a PF-phenomenon cannot straightforwardly solve the extraction problem. The problem is as follows. Extraction out of a CP which is in its 'base-position' derives an ill-formed output:

(17) (a) *Hans hat [daß Maria Peter geküßt hat] geglaubt.*
John has [that Mary Peter kissed has] believed

(b) **Wen$_i$ hat Hans [daß die Maria t$_i$ geküßt hat] geglaubt.*[8]
Who has John [that the Mary kissed has] believed

(c) *Wen$_i$ hat Hans geglaubt, [daß Maria t$_i$ geküßt hat].*
Who has John believed [that Mary kissed has]

The example in (17a) shows that extraposition is not in general obligatory. (17b-c) create a minimal pair: extraction out of the embedded CP is possible if it occurs in extraposed position (17c) but it is impossible if the CP occurs within the matrix clause.

It can be concluded that assuming extraposition to be a PF-phenomenon leads to wrong predictions even for the extraction problem. However there are more general problems for the PF-approach towards extraposition. This discussion will turn out to be a summary of arguments found in the literature. There has been a discussion within Transformational Grammar as to whether extraposition is a cyclic or a post-cyclic rule. Most arguments for the former view can be taken to argue against extraposition as a PF-rule. The simplest argument against extraposition as a PF rule is provided by cases that show that extraposition has an effect at LF. This is true both for the standard T-

[8] Notice that my analysis does not make any prediction about this phenomenon.

model and within the minimalist program. If movement takes place between SPELL-OUT (or S-structure) and PF it cannot have any influence at LF. However, Guerón (1980) gives evidence that extraposition can have influence at LF.

Quantifier Raising. Guerón (1980) gives some examples where it is shown that PP extraposition has consequences for quantifier raising (i.e. LF-movement):

(18) (a) *The owner of every car on the block will be fined*
(b) *The owner will be fined of every car on the block*

The sentence in (17a) is ambiguous between a narrow scope and a wide scope interpretation:

(19) (i) For all x, x a car, the owner of x will be fined
(ii) That individual x, such that x owns every car, x will be fined

According to Guerón (1980), the sentence in (17b) in which the PP containing the quantifier is extraposed exhibits just the wide scope reading. If extraposition would be a stylistic (PF) rule then both readings should be available in either case. Furthermore Guerón (1980) mentions in a footnote that if a complex NP exhibits just narrow scope reading then PP extraposition is forbidden:"

" Guerón (1980) gives another similar argument. This piece of evidence concerns the logical quantifier *only*. Extraposition has an effect on the acceptability of NPs containing *only*:
(i) (a) *Only those people who are interesting to talk to will be invited.*
(b) *Only those people will be invited who are interesting to talk to.*
(c) *The only man (there) who was interesting to talk to was invited*
(d) **The only man (there) was invited who was interesting to talk to.*
The same is true for the following examples containing *all* or a superlative respectively:
(ii) (a) *All those who could walk came*
(b) *All those came who could walk*
(c) *Those who could walk all came*
(d) **Those all came who could walk*
(iii) (a) *The friends are gone that I once had*
(b) **The best friend is gone that I ever had.*
We will return to *only* and superlatives in the following excursion. There we will see that in German extraposition out of these NPs does not result in ungrammaticality. The relevant NPs show an ambiguity which is lost if the constituent occurs in extraposed position. This fact however will be shown to derive from LCI. On basis of the present analysis Guerón's (1980) examples can no longer be interpreted as counter-evidence for treating extraposition as a PF-rule. The same is true for the following example due to Akmajan (1975). The complex NP in the following example is ambiguous:
(iv) *I have read a review of a book by three authors*
The PP *by three authors* can either modify the whole NP (*a review by three authors*) or only the

(20) (a) *The sum of these two numbers was requested*
(b) **The sum was requested of these two numbers*

Moreover Guerón (1980) shows that extraposition is input for the rules determining the logical scope of negation and polarity items:

(21) (a) **The names of any of those composers weren't called out yet.*
(b) *The names weren't called out yet of any of those composers.*
(22) (a) **Mark thinks that the extraposition transformation which has the slightest effect on LF hasn't been found yet.*
(b) *Mark thinks that the extraposition transformation hasn't been found yet which has the slightest effect on LF*

Finally an example due to Huybregts is mentioned showing that PP extraposition provides an input for the (S-structure) rules of Government and Binding:

(23) (a) *A man Bill knows arrived*
(b) **A man arrived Bill knows*
(c) *A picture of Mary was sent to her*
(d) **A picture was sent to her of Mary*

The latter example shows that coreference is not possible anymore if the PP is in extraposed position. This situation would not be expected if extraposition were a PF-rule.

Extraposition and wh-movement. Another argument against extraposition as a PF-rule is as follows: it can be shown that extraposition has influence on the acceptability of wh-extraction. Ross (1967) argues that the rule of extraposition must precede Question (i.e. wh-movement) because

> *"while no elements of subject clauses may be moved out of these clauses, by virtue of the Sentential Subject Constraint, if these clauses have been extraposed, elements in them become movable. This means that the rule of Extraposition affects the Rule of Question and therefore the former has to be ordered before the latter"* (Ross, 1967)

embedded NP *a book*. However, if the PP is extraposed it can just be interpreted as modifying the whole NP:
(v) *I have read a review of a book yesterday by three authors.*
This pattern similarly follows from our LCI. Under the present analysis it cannot be argued to be evidence against extraposition as a PF-phenomenon.

(24) (a) **The teacher who that the principal would fire was expected by the reporters is a crusty old battleax.*
(b) *The teacher who it was expected by the reporters that the principal would fire is a crusty old battleax.*

Notice that the same phenomenon is true for Norwegian where it is possible to extract out of an extraposed relative clause but this is not true for relatives which are in situ (i.e. adjacent to their head NP).

A similar effect was also seen in the German example (17) i.e. extraposition seems to be a necessary prerequisite for wh-extraction. Moreover, extraposition of the complement CP in German has an influence on the interpretation: the content of the complement CP must be presupposed if it occurs in non-extraposed position[10]. This is not true for extraposed complement CPs.

Subject-clause extraposition. Another argument concerns extraposition of subject-clauses:

(25) (a) *That Mary is a spy is well known*
(b) *It is well known that Mary is a spy*

If the subject clause occurs in extraposed position then 'it' has to be inserted in the subject position. It is not clear how this could be explained under the assumption that extraposition is a PF phenomenon. Notice that within the minimalist program it is assumed that after the branch to PF (SPELL-OUT) there is no access to the lexicon anymore.

Extraposition and binding. Another argument has to do with the distribution of anaphora. This argument is due to Jacobson & Neubauer (1976) who argue that extraposition is a cyclic rule. Consider the following paradigm:

(26) (a) *John thought that shaving himself would disturb Mary.*
(b) **John thought that it would disturb Mary to shave himself*
(c) *John thought that the fact that criticizing himself was hard disturbed Mary.*
(d) **John thought that it disturbed Mary that criticizing himself was hard.*

[10] cf. Trinker (in preparation) for this conclusion and an elaborate discussion of this phenomenon.

(e) *John thought that the fact that a picture of himself was hanging in the post office disturbed Mary.*
(f) **John thought that it disturbed Mary that a picture of himself was hanging in the post office.*

In all these examples a subject sentence contains an anaphora that is coreferent with an NP in the matrix clause. All of these sentences are grammatical if the subject clause is in SpecIP (or adjoined to IP depending on the analysis of subject sentences). However, if the subject clause occurs in extraposed position the sentences become ungrammatical. This situation is not expected under the assumption that extraposition applies at PF together with the assumption that binding applies at LF (which is a necessary assumption within the minimalist program). The same contrast is also found in other cases of extraposition, i.e. it is not a property which is related to subject-clauses:

(27) (a) *John thought that a picture of himself standing on the beach was given to Mary.*
(b) **John thought that a picture was given to Mary of himself standing on the beach*
(c) *%That a picture of himself standing on the beach was given to Mary surprised John.*[11]
(d) **That a picture was given to Mary of himself standing on the beach surprised John*
(e) *John thought that a book which showed a picture of himself was given to Mary.*
(f) **John thought that a book was given to Mary which showed a picture of himself.*
(g) *%That a book which showed a picture of himself was given to Mary upset John.*
(h) **That a book was given to Mary which showed a picture of himself upset John.*

Finally it has to be noticed that German NPs cannot extrapose. This fact cannot easily be explained under the assumption that extraposition is a PF phenomenon. Under the assumption that extraposition is an S-structure phenomenon this fact can be related with Case-assignment, i.e. it is only elements that do not need case that can be extraposed. Case is checked at LF. If extraposition would be a PF-phenomenon, then

[11] % indicates that speaker vary in their judgements but for the purpose to be shown here it is enough that there is a contrast to be found.

it should not interact with case theory. Under a PF-account something else has to be said about why it is just PPs and CPs that can occur in extraposed position.

I will take these arguments to crucially contradict the claim that extraposition is a PF-phenomenon. Given these facts it seems to me that Chomsky's (1995) remarks concerning optionality and the level where Topic-Focus has to be interpreted (as discussed in section II.1.1.) are problematic.

1.5. Conclusion

For reasons of space I will not discuss at any length other ways to derive the other paradigms (that are argued to be derived by PCI). Several more hierarchically-oriented possibilities could come to mind (like for example ECP- or cross-over violations respectively). However this seems to be the more standard way to treat such effects. Many developments within syntactic theory replace *precedence* by *c-command*, i.e. linear notions with hierarchical ones. Thus it seems somehow contra these major developments to reintroduce the notion of precedence. However, I would like to mention two crucial points in order to justify the analysis presented here. The first point concerns the fact that PCI provides a unified account for the paradigm discussed in the last sections. None of the structural accounts would achieve this result. In general this would not be evidence for a solution to be preferable to another one. However I think that there is a reason to still take it as an advantage. It concerns the grammaticality judgements the excluded sentences give rise to. All of them have a similar flavor. Simply speaking, the impression is that the identifyee is "the disturbing factor". Confronted with these data the relevant question to be asked is not: Why is phenomenon X not possible if the item Y is present? Rather what seems to be the adequate question is: Why is the item Y not 'licensed' once the phenomenon X takes place? The answer within the solution suggested here of course is the Novelty condition (resulting in PCI). This means that if in general the occurrence of Y is optional then process X can take place if Y is not present. However the solution just works if Y is also present at domain D. In interaction with the Optionality Hypothesis (introduced in section III.3.) we now make another prediction. All the paradigms above should be wellformed in case the pronominal is only syntactically licensed but can be expletive at domain D. This prediction is borne out as can be observed in the following examples:[12]

[12] Again the problem of left-adjunction does not arise here since adjunction to a pronominal element is in general excluded. Therefore the prediction is vacuous here.

(28) (a) *?Den Fleck zu entfernen ist es mir gelungen.*
The spot to remove is it me succeeded
(b) *?Den Fleck zu entfernen gelungen ist es mir gestern.*
The spot to remove suceed is it me yesterday
(c) *?Was ist es dir gelungen zu entfernen*
What is it you succeed to remove

Although the sentences are not perfectly well-formed there is a remarkable contrast w.r.t. the sentences in (6,11 and 13). As I have argued in section III.3. the pronominal can be expletive at domain D in case it is syntactically forced (it only needs to be licensed at one level of representation). Examples where the pronominal can be expletive at domain D (according to the Optionality Hypothesis) are much better than the ones where it needs to be licensed at domain D. The fact that the sentences are still not as good as we might expect could have to do with the fact that the preferred interpretation is still the one where the pronominal is present at domain D.

The second point which is in favor of the analysis suggested here is the following one. The present proposal simply states that the identifyee has to precede the identifier. From this (further reduced) constraint all the facts are derived. Proposals which make use of a structural account to derive the data discussed in the previous sections still have to include the stipulation that extraposition is obligatorily to the right. I have generalized and derived this stipulation. *Rightness* has been replaced by *precedence*. Of course, the two notions are related (as we discussed in section I.3.). Talking about left- or rightness implies talking about precedence. However the distinction between left- and rightness as used in other accounts is related to hierarchical structure there. In these analyses it has to be assumed that there is a distinction between a position to the left or to the right relative to another position. However, in terms of hierarchy left- or rightness should not make a difference. It should only be hierarchy that counts (i.e. whether an element X c-commands or is c-commanded by another element Y respectively). On the other hand, my precedence constraint is related to linear ordering. The notion of precedence IS obviously relevant for linear ordering. Moreover, I derive the precedence constraint from a principle which is independently motivated to be related to linear ordering, namely from the Novelty Condition. Therefore I conclude that structural accounts for the relevant paradigms all have to stipulate a rightness condition before they can even start to analyze them. In deriving this 'rightness condition' by means of the Precedence Constraint I do not need to explain these facts additionally.

Some of the paradigms discussed in the last section might indeed be syntactically wellformed but independently ruled out by another wellformedness condition (which holds at domain D), i.e. the precedence constraint. Some of the examples could additionally be ruled out by purely syntactic constraints. For example in some of the examples a Complex NP Constraint violation can still be at play in addition to the Novelty Condition violation. It seems to me that this is also a desirable result: for example the cases where it is extracted out of a relative clause seem much worse than extraction out of an extraposed clause correlated with a pronominal (where I do not think that the CNPC can apply).

2. Predictions of the Locality Constraint

Let me now turn to the last part of my proposal, namely to the Locality Constraint on Identification. The formulation of i-government is repeated below for convenience:

(29) Locality Constraint for Identification (LCI) (holds at S-structure)
The identifyee has to be i-governed by the Identifyee.

(30) i(dentificational)-government:
X i-governs Y if X c-commands Y (or X is a sister to Y) and there is no node Z, Z dominates Y and X c-commands Y.

If one follows Kayne (1993) in assuming that c-command is only visible for categories but not for segments, this definition results in the following domain (descriptively spoken):

(31) LCI: (informal version)
An Identifier can either be adjoined to the identifyee or to the first maximal projection dominating the identifyee.[13]

[13] In order to arrive at this descriptive result I have to assume that dominance is visible for segments. Otherwise the identifier could be adjoined to one maximal projection lower than the identifyee, since the node resulting of adjunction would not count as dominating the identifier, and therefore it could still c-command the identifyee.

2.1. Upward boundedness

LCI immediately predicts the well known fact that extraposition is *upward bounded* (cf. Ross 1967). In (32) below we find a topicalized CP that contains the head NP of a relative clause. The identifier must occur in a position where LCI can be fulfilled[14]

(32) (a) *$[_{CP2}[_{CP1}[_{CP}$Daß Peter sich auf das Fest gefreut hat] [das Maria veranstaltet hat], hat niemanden gewundert.*
That Peter REFL to the party looked forward has that Mary organized has, has nobody surprised
'That Peter was looking forward to the party, which Mary has organized surprised nobody'

(b) **$[_{CP2}[_{CP1}$Daß Peter sich auf das Fest gefreut hat] hat niemand gewundert] [das Maria veranstaltet hat]*

Moreover, it is a well known fact that upward boundedness can be violated if the nominal element construed with the extraposed clause is moved (cf. Ross 1967). The following example is due to Guerón & May (1984):

(33) *$[_S$Which spy e_i [does Angleton believe $[_S$ e_i that $[_S$ Burgess recruited e_i]]] who ultimately became a mole$_i$]]* (Guerón & May 1984: 16)

Nothing else has to be said about those facts. They immediately follow from the present analysis: it is predicted that the S-structure position of the nominal element in the matrix determines the adjunction site of the extraposed element. The identifier must be adjoined to the next maximal projection. Notice that within the present account upward boundedness of extraposition of constituents without a correlate in the matrix follows from the locality constraint on theta-Identification.[15]

[14] Interestingly, upward boundedness can be violated exactly in cases where one can argue that restructuring of a bisentential clause to a monosentential one has taken place:
(i) *??Peter hat [dem Mann den Computer zu bringen [den er nicht mehr braucht]] versprochen*
Peter has the man the computer to bring [which he not anymore needs] promised
'Peter promised to brig the man the computer, that he didn't need anymore.'
(ii) *Peter hat [dem Mann den Computer zu bringen] versprochen, den er nicht mehr braucht*
(i) shows that in case of restructuring extraposition to a position not violating upward boundedness results in a degraded output, (ii) involves an apparent violation of upward boundedness but the sentence is perfectly well-formed. This is expected under the assumption that we have reanalysis of the bisentential clause to a monosentential one (no matter which particular analysis for reconstruction one adopts).

[15] cf. also Wiltschko (1994) for a discussion.

2.2. Nesting

Another restriction for extraposition has descriptively been characterized as a nesting requirement on multiple extraposed constituents. This property has been analyzed in C&R (1990) as an interpretive requirement.[16] This requirement can be seen in the following examples:

(34) (a) *weil ein Mann$_i$ die Party$_j$ langweilig finden wird, die$_j$ Maria veranstaltet, der$_i$ nur Bier trinkt.*
because a man$_i$ the party$_j$ boring consider will, which$_j$ Mary organizes, who$_i$ only drinks beer
'because a man who only drinks beer will consider the party boring which Mary organizes'

(b) **weil ein Mann$_i$ die Party$_j$ langweilig finden wird, der$_i$ nur Bier trinkt, die$_j$ Maria veranstaltet.*

By the analysis I am presenting here, this requirement can be interpreted as a byproduct of the more general constraint on the possible adjunction site for extraposed elements which are construed with a correlate in the matrix (i.e. the LCI). The (S-structure) position of the identifyee determines the adjunction-site of the identifier. It follows that the higher the identifyee occurs in the matrix the higher the identifier must occur. Since we are dealing with right-adjunction the nesting requirement, which results in symmetry, follows. The higher an element is to the left of the tree the more left is it w.r.t. linear ordering. The higher an element is to the right of the tree, the more right it is w.r.t. linear ordering. Notice that with multiple right-adjoined elements the Locality Constraint is somehow visible. (In fact this is the only test for the relative height of the adjunction site of extraposed element which can rely solely on the surface occurrence of other elements). We do not have to assume an additional interpretive requirement like C&R (1990) have to. Their account does not immediately predict the nesting requirement since the correlate does not unambiguously determine the adjunction site for extraposition, i.e. the identifier can either be higher or lower as the

[16] The *nesting requirement* can be formulated as the following descriptive generalization: dependencies may not cross, but have to be nested. One way to analyze this generalization would be to state it as a restriction on chains (assuming an analysis which makes use of Pesetsky's (1982) Path Containment Condition (PCC)). However this approach seems to be problematic in various respects: first, not all possible chains obey nesting (for a discussion see Nakajima 1991). Secondly, if nesting (the PCC) were reduced to the ECP (i.e. to a Relativized Minimality approach à la Rizzi 1990) one would have to assume a movement approach for extraposition. Furthermore we would have to define extraposed elements as possible interveners for particular movement processes, which creates several problems, not to be discussed here.

correlate (cf. Appendix I. for a discussion). Moreover we also do not have to assume a specified adjunction site for extraposition.

The nesting requirement does not only hold for two relatives. As we predict with LCI it holds for all instances of Identification. This prediction is borne out as the following examples show:

(35) (a) *weil ein Mann$_i$ es$_j$ verhindert hat, [daß Maria das Fest besucht]$_j$, den$_i$ sie gar nicht kannte*
because a man$_i$ it$_j$ prevented has, [that Mary the party visits]$_j$, whom$_i$ she at all not knew
'...because a man whom she didn't know at all prevented Mary to go to the party'

(b) **weil ein Mann$_i$ es$_j$ verhindert hat den$_i$ Maria gar nicht kannte, [daß sie das Fest besucht]$_j$*

(36) (a) *weil ein Mann$_i$ daran$_j$ geglaubt hat [daß Maria das Fest besucht]$_j$, den$_i$ sie gar nicht kannte*
because a man$_i$ thereon$_j$ believed has [that Mary the party visits]$_j$, whom$_i$ she at all not knew
'...because a man whom she didn't know at all believed that Mary goes to the party'

(b) **weil ein Mann$_i$ daran$_j$ geglaubt hat den$_i$ sie gar nicht kannte, [daß Maria das Fest besucht]$_j$.*

(37) (a) *weil es$_j$ einen Mann$_i$ gestört hat den$_i$ sie gar nicht kannte, [daß Maria das Fest besucht hat]$_j$*
because it$_j$ a man$_i$ annoyed has whom$_i$ she at all not knew, [that Mary the party visited]$_j$
'...because it annoyed a man whom she didn't knew at all, that Mary visited the party

(b) **weil es$_j$ einen Mann$_i$ gestört hat [daß Maria das Fest besucht hat]$_j$, den$_i$ sie gar nicht kannte.*

(38) (a) *weil das Argument$_j$ einen Mann$_i$ aufgeregt hat der$_i$ das Fest besuchte$_j$, [daß Rauchen ungesund ist]$_j$.*
because the argument$_j$ a man$_i$ upset has who$_i$ the party visited [that smoking is unhealthy]$_j$

(b) **weil das Argument$_j$ einen Mann$_i$ aufgeregt hat [daß Rauchen ungesund ist]$_j$, der$_i$ das Fest besuchte*

Finally, the nesting requirement also supports the claim that the input for the constraint on the actual position of the identifying element is S-structure; consequently there are different possible positions for the extraposed relative depending on the actual position

of the antecedent. Such a view could not be taken under an analysis which makes crucial use of a specified adjunction site of extraposition in general (this is explicitly done e.g. in Ross 1967, Reinhart 1983, Müller 1994 and implicitly necessary in Guerón & May 1984 as well as C&R 1990; cf. Appendix I. for a discussion).

That S-structure is the input for the constraint can be seen in the following examples involving topicalization of the object to SpecCP: the ordering of the extraposed elements is the reverse of the one shown (37):[17]

(39) (a) *Einen Mann$_i$ hat es$_j$ gestört [daß Maria raucht]$_j$ [den$_i$ sie gar nicht kannte]$_i$.*
A man$_i$ has it$_j$ annoyed [that Mary smokes]$_j$ [who$_i$ she not knew]$_i$

(b) **Einen Mann$_i$ hat es$_j$ gestört [den$_i$ Maria gar nicht kannte]$_i$ [daß sie raucht]$_j$.*
A man$_i$ has it$_j$ annoyed [who$_i$ Mary not knew]$_i$ [that she smokes]$_j$

The same is true for any other reordering process of the antecedents in the matrix clause.[18]

We can also observe a restriction on the ordering of extraposed elements in case one of the extraposed elements does not have an antecedent in the matrix. As I briefly mentioned in section III.3. I am assuming that theta-identification is subject to a locality constraint as well (i.e. government). The result is then that those CPs can only be adjoined to VP (i.e. rather low in the tree):[19]

[17] This fact is also an important piece of evidence against approaches which state that the conditions on these cases of extraposition are constrained by a condition holding at LF (cf. Guerón 1980, Guerón & May 1984). If the relevant locality constraint (which is assumed to be the 'complement principle) did not hold before LF one would need to give an account for preventing 'LF-extraposition', i.e. moving the extraposed element to a position where it could fulfill the complement principle at LF. Thus LF-accounts still have to assume that it is the S-structure position which is the relevant input for the constraint imposed on the LF-condition.

[18] It has to be noticed that English seems to behave differently. At least for elements that have a certain scope domain (i.e. are moved at LF) it seems to be the case that the scope position (i.e. the LF-position) is the relevant input for the nesting requirement. This has been observed in Williams (1974). He argues that scope properties determine the adjunction site for extraposition. There is a distinction between scope determined by *precede* (which would result in the nesting effect) and scope determined by something else. If an element has S' scope, then an associated extraposed clause must necessarily adjoin to S' as well, the S-structure position apparently does not play a role. I have nothing to say about these facts about English.

[19] I have coindexed the verb and its complement to indicate the relation of θ-identification.

(40) (a) *Ein Mann$_i$ hat Maria gebeten$_j$ [das Buch zu lesen]$_j$ [den$_i$ sie noch nie gesehen hat]$_i$*
A man$_i$ has Mary asked$_j$ [the book to read]$_j$ [whom$_i$ she yet never seen has]$_i$
'A man asked Mary to read the book whom she has never seen before.'

(b) **Ein Mann$_i$ hat Maria gebeten$_j$ [den$_i$ sie noch nie gesehen hat]$_i$, [das Buch zu lesen]$_j$*
A man$_i$ has Mary asked$_j$ [whom$_i$ she yet never seen has]$_i$[the book to read]$_j$

However, things get more complicated if we look at the following paradigm, which suggests that the ordering of multiple extraposed elements is optional:

(41) (a) *Peter hat eine Frau$_i$ gebeten$_j$ [das Buch zu lesen]$_j$ [die$_i$ er gar nicht kannte]$_i$*
Peter has a woman$_i$ asked$_j$ [the book to read]$_j$ [which$_i$ he not at all knew]$_i$
'Peter has asked a woman he didn't know to read the book.'

(b) *Peter hat eine Frau$_i$ gebeten$_j$ [die$_i$ er gar nicht kannte]$_i$ [das Buch zu lesen]$_j$*

It can be shown that in case we find an apparent option w.r.t. linear ordering we are dealing with two different positions of the head NP of the relative. This can be shown by means of the position of adverbs marking the VP boundary[20]:

(42) (a) *?*Peter hat [$_{VP}$ja doch eine Frau$_i$ gebeten$_j$ [das Buch zu lesen]$_j$ [die$_i$ er gar nicht kannte]$_i$*
Peter has Prt. a woman asked [the book to read]$_j$ [who$_i$ he didn't know]$_i$

(b) *Peter hat eine Frau$_i$ [$_{VP}$ja doch gebeten$_j$] [das Buch zu lesen]$_j$ [die$_i$ er gar nicht kannte]$_i$*

(c) *Peter hat [$_{VP}$ja doch eine Frau$_i$ gebeten$_j$ [die$_i$ er gar nicht kannte]$_i$ [das Buch zu lesen]$_j$*

(d) *?*Peter hat eine Frau$_i$ [$_{VP}$ja doch gebeten$_j$ [die$_i$ er gar nicht kannte]$_i$ [das Buch zu lesen]$_j$*

This means that the apparent optionality of the relative ordering can be reduced to two different S-structure positions of the relevant NPs. Therefore to two different possible positions of the extraposed relative clause depend on the actual position of the

[20] For the purpose of the present argumentation I will simply abstract away from the problems induced by the assumption that *ja doch* really marks the VP-boundary.

antecedent. Again such a view could not be taken under an analysis which makes crucial use of a specified adjunction site of extraposition.

2.3. Split Antecedents

As I have already mentioned in section IV.5. there is a difference in grammaticality of split antecedents depending on the position of the two relevant antecedents. The observation is that split antecedents are possible if they occur in coordinated clauses whereas it is not possible if they occur within one single clause:

(43) *[A man]$_i$ entered the room and [a woman]$_j$ went out [[who]$_{i,j}$ were quite similar]$_{i,j}$.*

(44) **[A man]$_i$ met [a woman]$_j$ [who$_{i,j}$ were quite similar]$_{i,j}$.*

In case of coordinated clauses (43) the relative clause is related to two antecedents which are syntactically in the same position (i.e. both DPs occupy the subject position of the coordinated IP). On the other hand the sentence in (44) involves a single sentence. The two antecedents occupy different syntactic positions (one is in the subject the second one in the object position). The reason for the ungrammaticality of (44) as opposed to the wellformedness of (44) can be explained by LCI. In (43) an appropriate adjunction site (satisfying the condition for both head NPs) is available. No such adjunction site is available in case of (44). Since there are two different positions of the antecedents, the identifier would have to occur in two different positions to fulfill LCI. In either of the two positions available LCI would be violated w.r.t. the other identifyee. I will come back to some more instances of split antecedents embedded in coordinated constituents in Appendix II.

2.4. Stress and LCI

There is another interesting paradigm concerning the LCI on the one hand as well as the previously discussed phenomenon of stressed determiners on the other hand. I will not provide a detailed (technical) solution for the following facts but only describe them. The actual implementation depends on an appropriate theory of focus. The relevant facts are as follows: If the identifyee is more deeply embedded in another constituent there is the possibility of 'constituent-internal-extraposition', i.e. the identifier can occur in a position which is not adjacent to the identifyee but still not in the position following the verb. It is possible to adjoin the identifier to a constituent

that is not itself the identifier but also not a projection of the main projection line. As expected, (under neutral intonation) it is not possible to extrapose the identifier to a position which is outside the i-government domain. However, stress on the determiner extends this local domain. In all the relevant cases extraposition to a position following the verb in final position is only possible if the determiner is stressed.[21]

2.4.1. Identifyees embedded within DP

In the following examples we find the identifyee realized as a possessor (occupying SpecDP):

(45) (a) $[_{DP}$ [DP]$_i$ D... $]$ CP$_i$ (b) *...$[_{DP}$ [DP]$_i$ D...$]$.....V$]$ CP$_i$

(46) (a) *Peter hat [dem Mann$_i$ [den er kannte]$_j$ sein Fahrrad$_j$] repariert*
Peter has [the man$_i$ [who he knew]$_j$ his bike] fixed.
'Peter has fixed the man he knows' bike'

(b) *Peter hat [dem Mann$_i$ sein Fahrrad$_j$] [den$_j$ er kannte]$_i$ repariert*
Peter has the man$_i$ his bike$_j$ [who$_j$ he knew]$_i$ fixed

(c) **Peter hat [dem Mann$_i$ sein Fahrrad$_j$] repariert [den$_j$ er kannte]$_i$*
Peter has [the man$_i$ his bike$_j$] fixed [who$_j$ he knew]$_i$

(d) *Peter hat [DEM Mann$_i$ sein Fahrrad$_j$] repariert [den$_j$ er kannte]$_i$*
Peter has [THE man$_i$ his bike$_j$] fixed [who$_j$ he knew]$_i$

Here the relative clause can either be adjoined to the identifyee (46a) or to the dominating DP (46b) obeying LCI. Furthermore LCI predicts that it is not possible to adjoin the identifier to any other node (46c). However, as example (46d) shows, this is not really true: the identifier can also occur in extraposed position (following the verb), but this is only possible if the determiner is stressed. Therefore we can conclude that the locality domain can be extended if the determiner is stressed. The same phenomenon can be observed in the following paradigm:

(47) (a) $[_{DP}$ D...[DP]$_i$$]$, CP$_{i/i}$ (b) ...$[_{DP}$ D...[DP]$_i$$]$,... V$]$ CP$_{i/*_i}$

(48) (a) *[Einer$_i$ [der Männer]]$_j$ der$_j$ gerne Bier trinkt ist gekommen.*
[One$_i$ [the$_{GEN}$ men]]$_j$ who$_{i,j}$ likes beer drinking is come
'One of the men who likes to drink beer came'

[21] Most of the data are drawn from Wiltschko (1994), where more examples concerning this phenomenon are presented.

(b) *[Einer [der Männer]$_i$] [die$_j$ gerne Bier trinken]$_j$ ist gekommen.*
[One [the$_{GEN}$ men]$_i$ [who$_{pl\,j}$ like beer drinking]$_j$ is come

(c) *[Einer$_i$ [der Männer]]$_i$ ist gekommen [der$_i$ gerne Bier trinkt]$_i$*
[One$_i$ [the$_{GEN}$ men]]$_i$ is come [who$_{s\,i}$ likes beer drinking]
'One of the men came who likes to drink beer'

(d) **[Einer$_i$ [der Männer]$_j$]$_i$ ist gekommen [die$_j$ gerne Bier trinken]$_j$*
[One$_i$ [the$_{GEN}$ men]$_j$]$_i$ is come [who$_{pl\,j}$ likes beer drinking]$_j$
'One of the men came who likes to drink beer'

(e) *[Einer$_i$ DER Männer]$_i$ ist gekommen [die$_j$ gerne Bier trinken$_{pl}$*
[One$_i$ [THE$_{GEN}$ men]]$_i$ is come [who$_j$ likes beer drinking]$_j$

(f) **[Einer$_i$ DER Männer]$_i$ ist gekommen [der$_i$ gerne Bier trinkt]$_i$*
[One$_i$ [THE$_{GEN}$ men]]$_i$ is come [who$_i$ likes beer drinking]$_i$

The paradigm is as follows. In case the identifier occurs adjacent to the DP *einer der Männer*, then (in accordance with LCI) it can be construed with either with the whole DP *einer der Männer* (48a) or only with the embedded one *der Männer* (48b) . The two cases differ as to which relative pronoun is used, in the former case it must be singular, in the latter it occurs in plural. Notice that in this case (i.e. in order to identify the embedded DP) the identifier can either be adjoined to the embedded DP itself or to the higher DP (the difference here is not detectable on basis of word order here). If however the relative clause occurs in extraposed position (i.e. in a position following the verb), then it can just be construed with the whole DP but not with the embedded one (as predicted by LCI) (48c/d). Again the situation changes in case the determiner of the embedded DP is stressed. In this case the relative clause can (and in fact must) be construed with the embedded DP (48e/f).

2.4.2. Identifyees embedded within (topicalized) VP.

The same phenomenon can also be observed in case an identifier is dependent on a constituent embedded within a topicalized VP.

(49) (a) $[_{CP}[_{VP}...DP_i...V]\ CP_i]\ [_{C'}.....t$
(b) $*[_{CP}[_{VP}...DP_i...V].....t\ ]\ CP_i$

(50) (a) *$[_{CP}[_{VP}$ [Den Mann]$_i$ gesehen [der$_i$ Bier trinkt]$_i$] hat Peter gestern auf der Party]*
[[The man]$_i$ seen [who$_i$ beer drinks]$_i$] has Peter yesterday at the party

(b) *[[DEN/*den Mann]$_i$ gesehen] hat Peter gestern auf der Party [der$_i$ Bier trinkt]$_i$*
[[THE/*the man]$_i$ seen] has Peter yesterday at the party [who$_i$ beer drinks]$_i$

The first maximal projection dominating the identifyee is VP. Therefore the identifier can move along with the topicalized VP (it is adjoined to VP). If the identifier occurs in a position following the verb the present analysis forces us to assume that it is adjoined to CP.[22] Remember, that it is the S-structure position of the identifier that determines the adjunction site of the identifier. As expected, the relative clause can only occur in this position in case the determiner is stressed since LCI is violated (CP is not the first maximal projection dominating the identifyee).

Notice that assuming CP as the adjunction site of the identifier predicts the following pattern w.r.t. multiple extraposed relative clauses. We expect the reverse relative order of the extraposed constituents than the one we found in case of examples without VP-topicalization. (cf. section VI.2.2.). This prediction is indeed borne out:

(51) (a) *Der Mann$_i$ hat das Bier$_j$ getrunken, [das$_j$ Maria ihm empfohlen hat]$_j$ [der sonst nur Wein trinkt]$_i$*
the man$_i$ has the beer$_j$ drunk [that Mary him recommended has]$_j$ [who otherwise only wine drinks]$_i$

(b) **Der Mann$_j$ hat das Bier$_i$ getrunken, [der$_j$ sonst nur Wein trinkt]$_j$ [das$_i$ Maria ihm empfohlen hat]$_i$.*
the man$_j$ has the beer$_i$ drunk [who otherwise only wine drinks]$_j$ [that Mary him recommended has]$_i$

(52) (a) *[DAS Bier$_i$ getrunken] hat der Mann$_j$ gestern [der$_j$ sonst nur Wein trinkt]$_j$ [das$_i$ ihm Maria empfohlen hat]$_i$.*
[The beer$_i$ drunk] has the man$_j$ yesterday [who$_j$ otherwise only wine drinks] [which$_i$ him Mary recommended has]

(b) **[DAS Bier$_i$ getrunken] hat der Mann$_j$ gestern [das$_i$ ihm Maria empfohlen hat]$_i$ [der$_j$ sonst nur Wein trinkt]$_j$.*
[The beer$_i$ drunk] has the man$_j$ yesterday [which$_i$ him Mary recommended has] [who$_j$ otherwise only wine drinks]

[22] Notice that another possibility (to be found in the literature) is that the identifier is still adjoined to VP. In this case VP-topicalization would only move the lower node of the two nodes created through adjunction. (cf. C&R for such a view)

In (51) the VP is in its base position and the extraposed relative clauses obey the 'nesting requirement'. This results in a linear ordering where the subject relative clause follows the object relative clause. However, if the VP is topicalized the relative order of the two relative clauses changes (52) as expected under the present analysis: again it is the S-structure position of the identifyee that determines the adjunction site of the identifier. Notice that this result is not expected under an analysis which makes use of a specified node for extraposition in general (i.e. if it is assumed that extraposition invariantly targets VP).

2.4.3. Identification of NP (X) is not possible if the identifier is in extraposed position.

In this section I will discuss how LCI interacts with the presuppositions discussed in section III.4. As I have argued there, the identifier can either identify NP (X = the set of individuals out of which the actual discourse referent is taken) or DP (x = the actual/intended referent). Consider the following example:

(53) *Peter hat gestern [das neue [Argument]$_X$], [daß Maria gerne Bier trinkt]$_{X_x}$ präsentiert*
'Peter has presented [the new [argument]$_X$]$_x$ [that Mary likes to drink beer]$_{X_x}$'

The denoted discourse referent *the argument* has two distinguishing properties P_1 and P_2. It is *new* (= P_1) and it has the content *that Mary likes to drink beer* (= P_2). Example (53) gives rise to an ambiguity. The two possible readings are given in (54) and (55) i.e. there are two different presuppositions possible:

(54) Assertion: Peter presented P_1 NP_i P_2
Presupposition: There is some NP_j P_2 (NP_j is not NP_i) such that $\neg P_1$ is true of NP_j, i.e. [[P_2 NP_j] $\neg P_1$]
Paraphrase: There are other arguments that Mary drinks beer, but they are not new

(55) Assertion: Peter presented P_1 NP_i P_2
Presupposition: There is some P_1 NP_j (NP_j is not NP_i) such that $\neg P_2$ is true of NP_j, i.e. [[P_1 NP_j] $\neg P_2$]
Paraphrase: There are other new arguments, but they do not say that Mary drinks beer

This ambiguity arises because it can either be P_1 or P_2 that counts as the distinguishing property for picking out the relevant discourse referent. The second property is then interpreted as a common property of the set the individual is picked out. Although it restricts the set of *possible arguments* (therefore creating a subset) it is still not enough to pick out one particular discourse referent. I assume that the first reading (54) corresponds to identifying X and the second reading (55) corresponds to identifying x.[23] LCI predicts that an identifier in extraposed position can only identify x (the DP) but not X (the NP). If it occurs in extraposed position (following the verb), the DP is an intervening maximal projection for Identification of X (NP). Therefore X cannot be identified any longer - the sentence is disambiguated. This is a correct prediction. The following sentence can just have the interpretation in (55), i.e. it is just x that can be identified:

(56) *Peter hat gestern das neue Argument präsentiert, daß Maria gerne Bier trinkt.*
'Peter has presented the new argument yesterday, that Mary likes to drink beer

Again, it can be observed that stress on N extends the local domain for Identification. Both readings given above are available, even if the identifier occurs in extraposed position:

(57) *Peter hat gestern das neue ARGUMENT präsentiert, daß Maria gerne Bier trinkt.*
Peter has yesterday the new argument presented, that Mary likes beer drinking

As I have argued in section IV.7.2., focus on N results in the necessity of associating the identifier with NP rather than DP. (57) indicates that it does not even matter that LCI is violated. We can therefore conclude that it is a general property of stress to extend the locality restriction (no matter whether N or D is stressed).

As already mentioned, I will not go into possible analyses for the phenomenon that local domains can be extended by means of focus. Suffice it to say that it is not an unexpected phenomenon: focus CAN generally extend local domains as I will discuss immediately. In most analyses concerning focus this phenomenon is captured by assuming that focus induces movement at LF. This is however not a real possibility for the present analysis, since we have evidence that it is the S-structure position of the

[23] Notice that this is a simplified assumption. However, I will show in the following excursion on *einzig-NPs* that this assumption is legitimate. It amounts to saying that it is either the set N alone which has the restriction denoted in the identifier, or A and N provide the set out of which the actual individual is taken out.

identifier that determines the adjunction site.[24] It seems that we have to make use of a percolation approach to assign the scope of focus.

Notice that there is independent evidence for the assumption that focus can extend a given locality domain. It is a standard assumption that a focus particle like *sogar* has to c-command its associate (cf. Bayer 1991). However, if the associate is stressed, then it can occur in a position that is not in the c-command domain of the focus particle:

(58) (a) *Maria hat gestern nachmittag sogar den Hans getroffen*
Mary has yesterday afternoon even the Hans met
(b) *Maria hat sogar gestern nachmittag den Hans getroffen*
Mary has even yesterday afternoon the Hans met
(c) *Den Hans hat die Maria sogar getroffen*
The Hans has the Mary even met
(d) *Den HANS hat die Maria sogar getroffen*
The HANS has the Mary even met

Both in (58a) and (58b) *sogar* can associate with *Hans*. This is however not a possible option for (58c). Here *Hans* is not in the c-command domain of the focus particle. *sogar* necessarily associates with the verb. However, if *Hans* is stressed, then it can again be associated with the focus particle. On basis of this paradigm we can again conclude that intonation can provide a sufficient condition for association with a certain element, even if the default locality constraint is violated. The same phenomenon is found in English as well. Karttunen & Peters (1979) observe that under 'normal' circumstances the focus particle 'even' is associated with its sister. But they conclude that:

"In spoken English, the intended focus of even *can be marked by stress to reduce ambiguity."*

(59) (a) *BILL, even likes Mary*
(b) *Mary even wants to go out with BILL*
(c) John even talked about NIXON in his commencement address
(Karttunen & Peters, 1979: 24)

[24] The main evidence for this claim concerns nesting effect (cf. section VI.2.2. and VI.2.4.2.). We will see some more evidence for S-structure as the determining level in chapter VIII. and Appendix I.

Whatever reason is responsible for this phenomenon, it seems to be at play in the examples we are considering here as well. Notice however that there is an example given in Karttunen & Peters (1979) which suggests that it is not LF-movement (similar to QR) that is responsible for the scope assignment of focus. The example is as follows:

(60) *It is hard for me to believe that Bill can understand even SYNTACTIC STRUCTURE.*

Karttunen & Peters (1979) distinguish between the focus and the scope of *even*. Their notions correspond to the *semantic value* and the *focus semantic value* in Rooth's terms. In my terms the 'scope' of *even* corresponds to the property of the *superset*, i.e. the presupposition we arrive at. The sentence above is ambiguous w.r.t. the scope of *even*. This results in two different presuppositions.[25] Under one reading there is a presupposition that there are other books about which it is hard for me to believe that Bill can understand them. This would correspond to a wide scope-reading of *even*. The other (narrow-scope) reading induces the presupposition that there are other books that Bill can understand besides *Syntactic Structures*. It seems to me that the possibility to assign *even* wide scope indicates that it is not just a matter of LF-raising, since other instances of LF-movement (i.e. QR) are strictly clause bound. This fact might be an indication that focus-(scope) assignment is NOT to be derived by means of movement at LF, at least not in a very straightforward way.

3. There is no specified adjunction site

There is another result to be derived from the present analysis. Notice that I am not assuming a specified node of adjunction for extraposed constituents. The adjunction site is simply derived by means of LCI. If there is no identifyee then the present analysis makes two predictions. First PCI is not expected to be at play, in other words there should be an option as to whether the relevant constituent is realized to the left or to the right of the main projection line. Second there should not be any determined position for 'extraposition' - the position should be quite free. Let us first turn to the second prediction.

[25] It has to be noted that Karttunen & Peters (1979) talk about *conventional Implicature*, i.e. they redefine several cases of *pragmatic presupposition* as being instances of a conventional Implicature. I will not go into this discussion and stick to the term presupposition.

3.1. The Scope of left-adjoined adverbials

There is no way to determine the position of right adjoined elements in terms of word order: right-adjunction to any node of the clause (VP,IP or CP) results in a word-order position following the verb. But there is still an indirect way to determine at least the relative height of the adjunction site for extraposed elements, and it has to do with the scope of adverbials.

Adverbs which are adjoined to the left of the sentence take scope over the constituent they are adjoined to (where default scope assignment depends on sisterhood). So in German we find a difference in the following examples:

(61) (a) *Peter wird $[_{XP}$ $[_{ADV}$so schnell er kann$]$ $[_{XP}$Maria die Geschichte erzählen$]$*
Peter will $[_{XP}$ $[_{ADV}$as fast as he can$]$ $[_{XP}$ Mary the story tell$]$
'Peter will tell the story to Mary as fast as he can'

(b) *Peter wird $[_{XP}$Maria die Geschichte $[_{YP}$ $[_{ADV}$so schnell er kann$]$ $[_{YP}$ erzählen$]$*

The corresponding paraphrases for the different interpretations are given below:[26]

(62) (a) Adverb is adjoined higher (to XP)
As fast as he can from the time of utterance. Peter will start to tell Mary the story.

(b) Adverb is adjoined lower (to YP)
The action of telling will be as fast as it is possible for Peter

3.2. Ambiguity with right-adjoined adverbials

As we expect under the assumption that these adverbials can be right-adjoined to any constituent (they can be left adjoined to) we find both interpretations equally available in case the adverb is right-adjoined. There is no preferred interpretation as in the sentences where the adverb is left-adjoined to different nodes (resulting in different word-order). The sentence in (63) can have both interpretations given in (62):

[26] It has to be noted that with neutral intonation there is a strong preference for the reading corresponding to the relevant paraphrases given above. However, as expected (given the remarks concerning intonation as a device to 'violate' unmarked locality domains) stress on the relevant (associated) element makes the relevant reading available even if it is not the sister of the adverbial.

(63) *Peter wird Maria die Geschichte erzählen, so schnell er kann*
Peter will Mary the story tell, as fast he can

3.3. Scope properties of elements adjoined to the right mirror scope properties of elements adjoined to the left

The data under consideration provide also an argument for the symmetric nature of hierarchical structure: If we look at examples with an extraposed adverbial and an extraposed argument clause, we find scope properties that mirror the scope properties of left-adjoined elements:

(64) (a) *Peter wird Maria erzählen [daß Hans ein Spion ist] [so schnell er kann]*
Peter will Mary tell [that John is a spy] [as fast as he can]
'Peter will tell Mary that John is a spy, as fast as he can'

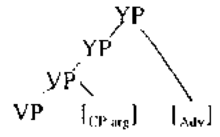
(b) *Peter wird Maria erzählen [so schnell er kann] [daß Hans ein Spion ist]*

The sentence in (64a), where the (right-adjoined) argument CP precedes the (right-adjoined) adverbial and therefore it must be adjoined lower, corresponds to the paraphrase in (65a). On the other hand the sentence in (64b), where the adverbial is lower than the argument CP, corresponds to the paraphrase in (65b):

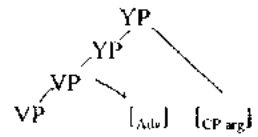
(65) (a) As fast as possible from the time of utterance Peter will tell Mary that John is a spy.

(b) The action of telling will be as fast as possible.

(66) (a)



(b)



It seems to be the case that the position to the right of the sentence (here made visible through another extraposed element) mirrors the scope properties of the position to the left of the verb. The scope properties clearly show the symmetry between left- and right-adjunction: for left-adjoined elements it is the case that they precede the elements they take scope over. For right adjoined elements it is the other way round: they follow

the elements they take scope over.[27]

These paradigms confirm both predictions mentioned at the beginning of this section.
1) PCI does not apply: once adjoined to a given node the relevant constituent can either be linearized to the 'left' or to the 'right' of the node they are adjoined to.
2) LCI does not apply: The elements under consideration can adjoin to any constituent they can take scope over.

Let me provide some more examples for the prediction in 1). In all of the following cases there is no (overt) identifyee present.

(67) (a) *Peter hat gesagt [daß Maria gerne Bier trinkt]*
Peter has said [that Mary likes beer drinking]
(b) *Peter hat [daß Maria gerne Bier trinkt] gesagt.*
Peter has [that Mary likes beer drinking] said

(68) (a) *Peter hat ein Bier getrunken [nachdem er Maria getroffen hat]*
Peter has a beer drunk [after he Mary met has]
(b) *Peter hat [nachdem er Maria getroffen hat] ein Bier getroffen.*
Peter has [after he Mary met has] a beer drunk

(69) (a) *Peter hat ein Bier getrunken [weil er Maria getroffen hat]*
Peter has a beer drunk [because he Mary met has]
(b) *Peter hat [weil er Maria getroffen hat] ein Bier getroffen.*
Peter has [because he Mary met has] a beer drunk

These data again indicate that it is in principle possible to either left- or right-adjoin. This is expected since left and right should not make a difference. However if some other factor (i.e. the Novelty Condition) is intervening then left-adjunction results in an ill-formed output.

4. Extraposition is optional

Obviously the previous discussion amounts to saying that *extraposition is optional*. The optionality is twofold. First, for extraposition without an identifier it is linearization that seems to be optional. Secondly, if an identifyee is present the LCI results in an option: i.e. there are in principle two possible adjunction sites (one resulting in a

[27] Notice that this piece of evidence is a crucial argument for assuming right-adjunction. Assuming a right-branching tree would make wrong predictions concerning these scope properties. Therefore these facts can be taken as an argument for symmetry in hierarchical structure.

position adjacent to the identifyee and one resulting in a discontinuous serialization.) Given that any syntactic difference should induce some interpretational difference (cf. the Optionality hypothesis discussed in section III.3.) we have to address this question here, too. We have to ask whether there is any difference in meaning corresponding to the different positions. In the present section I will show that there is indeed such a difference. However, I will just mention some possibilities to analyze this difference.

First, consider the cases where LCI allows for two different adjunction sites. Thus let us look at relative clauses in extraposed and non-extraposed position respectively:

(70) (a) *Maria hat [den Mann], [der gerne Bier trinkt], gesehen*
Mary has [the man], [who likes beer drinking], seen
(b) *Maria hat [den Mann], gesehen [der gerne Bier trinkt],*
Mary has [the man], seen [who likes beer drinking],

It has been noticed (for example in Beneš 1968) that the difference between the two serializations above corresponds to a difference in the communicative value:

> *"Die Satzkomponente, die die Geltung eines Satzes oder eines Satzäquivalents hat, wird dann ausgeklammert, wenn sie im Satzganzen als zweites (eigenes oder neues) Mitteilungs- und Intonationszentrum hervorgehoben werden soll"*
> (Beneš 1968/79: 328)
> (The clausal constituent is extraposed if it should be marked as a second (or new) information- or intonation center within the whole sentence. [translation MW])

The examples in (70) differ in the following way. In (70b) the relevant information is the fact that it was *the man who likes to drink beer* that was seen by Mary. On the other hand in (70a) the relevant information is more related to the fact that *Mary has seen this man*. The difference here again has to do with discourse (or information convey) rather than being truth-conditionally relevant. There are in principle two different possibilities to interpret this phenomenon. It could either be a matter of hierarchical configuration. Here it would be relevant that in case of extraposition the verb is within the scope of the relative clause, otherwise it is not. The other possibility would be to relate this fact to linear word-order. Here the sentences in (70) differ as to whether the relative clause precedes or follows the verb.

The second apparent optionality corresponding to an interpretational difference concerns sentences without an identifyee. Consider the following:

(71) (a) *Hans hat [daß Peter ein Spion ist] gesagt*
John has [that Peter a spy is] said
(b) *Hans hat gesagt [daß Peter ein Spion ist]*
John has said [that Peter a spy is]

Abstracting away from the fact that sentence (71a) is slightly degraded, there IS an interpretational difference to be observed. If the complement clause occurs in non-extraposed position it receives an interpretation that is reminiscent of the interpretation we arrive at in case a pronominal is present in the matrix. Given our analysis the presence of the pronominal results in the *familiarity*-presupposition. The following context is appropriate to utter a sentence like (61a):

(72) S1: *Hans hat geglaubt, daß Peter ein Spion ist.*
Hans has believed that Peter a spy is
S2: *Nein du irrst dich. Hans hat [daß Peter ein Spion ist] GESAGT, nicht nur geglaubt.*
No you err REFL. H. has that Peter a spy is SAID not only believed
'No you are wrong. John has said that Peter is a spy. He didn't only believe it.'

The difference seems to be related to informational structure: It is either the content of the complement clause, or the matrix verb which constitutes the information center. Notice that in this case the hierarchical configuration is the same in both sentences. They differ in word-order. Moreover, here we find a further difference. The sentences in (72a) and (72b) differ as to how the θ-role is assigned (making use of Higginbotham's (1985) assumption concerning different ways of θ-assignment). In non-extraposed position the CP can be directly (autonomously) θ-marked, whereas in non-extraposed position θ-assignment has to be mediated by means of θ-Identification. Notice that Higginbotham (1985) argues for a difference in interpretation corresponding to the two ways of θ-assignment in case of adjectival modification.

For the present discussion it should suffice that we have shown that the syntactic option indeed corresponds to a difference in meaning. However, I leave open the question as to how this difference should be analyzed.

EXCURSION
Applying the Analysis: A Case Study on *einzig*-NPs

1. Introduction

In previous chapters we have seen that an identifier can either be associated with x (i.e. D) or with X (i.e. N), respectively. Several different considerations have lead to this conclusion. Several problems were left open there especially concerning the difference between Identification of X and x. The present chapter is intended to give a more detailed discussion concerning this difference. We will look into some more data involving different DPs. In order to doing so in a precise way I will limit myself to a very restricted empirical domain. I have chosen an instance where the different readings can be seen more easily than in previous examples, namely DPs containing *einzig* ('only'). Since uniqueness is associated with its meaning it is not surprising that examples with *einzig* are a good empirical domain to study Identification. Remember that we have said that in case of Identification uniqueness and discourse availability occur separately whereas in other cases they go together. The identifyee was analyzed as introducing uniqueness whereas the identifier makes the discourse referent available. Consider under this light the following NP:

(1) *der einzige Mann, der an diesem Rennen teilnimmt*
the only man who at this race participates
'the only man that participates the race'

This NP allows for two different readings that can roughly be paraphrased in the following way:

(2) (a) of the set of men there is just one who participates in the race
(b) of the set of individuals participating in the race there is just one who is a man

These paraphrases show that the ambiguity concerns the presupposed set of individuals we talk about. It is a matter of the appropriateness of a given sentence in a given context. It might be argued that we do not have to encode this difference in meaning in Grammar. However, in the course of this chapter we will see that certain syntactic configurations exclude one of the two readings. Since it can be shown that different syntactic environments can influence the possible readings we can conclude that Grammar has to encode the different interpretations. I will again assume that the

relevant level for doing so is domain D.

2. The interpretation of *einzig*-NP'

How do the two different interpretations arise? As I have already said: *einzig* is associated with a uniqueness interpretation that can be paraphrased as follows:

(3) There is a (presupposed) set of individuals and one member of these individuals is distinguished from the others by being the only one who fulfills the property P.

einzig overtly expresses what we have assumed for Identification in general: there must be a distinguishing property of the denoted individual which suffices to pick this individual out of a given set.[1] As the two paraphrases in (2) indicate the difference in interpretation arises because there are two properties available. One of these properties can be the distinguishing property for the individual to be picked out. One of the properties is provided by N and the other one by the relative clause. Both of them can either be the distinguishing property of the individual or the property of the set out of which this individual is picked. In the case above this amounts to saying that we are talking about *a set of men, and one men is distinguished from the others by the property of participating the race*. The other reading is obtained by interpreting the property of *being a man* as the distinguishing property. If something is interpreted as a distinguishing property there must necessarily be some set out of which the denoted individual is picked. This set must be defined as having a particular property in common. The only other property available is the one denoted by the relative clause. Thus we get the interpretation that *out of the set of individuals participating the race there was just one man*. This is exactly the interpretation we argued to be an instance of Identification of X - since it is the whole set out of which the denoted individual is taken which is interpreted as having the property denoted in the relative clause.

In short we can again say that the ambiguity arises by either identifying x or X. *einzig* is thus somehow similar to focus, which as we have already seen can either be

[1] Notice that the presupposition above is reminiscent of Bach's generalization, the only difference being that using *einzig* exhausts the set. In case of relative clauses we get the result that there is at least one other individual who does not have the property denoted by the identifier. In case of einzig we get the presupposition that all other individuals in the relevant set do not have the property denoted by the identifier. Thus the relevant examples are similar to what we found in case of stress on the determiner.

associated with the determiner or the noun itself. In fact we can assume that *einzig* is a focus particle. So the presence of *einzig* has the same effect as stress (by means of intonation). Another possibility for analyzing the ambiguity would be to claim that *einzig* carries quantificational force inducing a tripartite structure as indicated below:

(4) Opx $|_{RC}$....x....$|$ $|_{NS}$....x....$|$

Elements in the restrictive clause restrict the range over which the operator functions. Assuming this analysis for operator-variable constructions the ambiguity induced by (1) could then be derived by the assumption that it can either be the NP or the restrictive relative clause that is mapped into the nuclear scope. If the NP is mapped into the nuclear scope the relative is mapped into the restrictive clause resulting in the reading where we talk about the individuals participating the race. If mapping takes place the other way round, i.e. the NP is mapped into the restrictive clause and the relative in the nuclear scope we derive the reading where we talk about a set of men. I think that the possibility to translate the present analysis into the 'tripartite-structure-analysis' (à la Heim 1982) sheds some light on the mapping hypothesis suggested in Heim (1982) as well as Diesing (1992). In the course of this section I will argue that my analysis has empirical as well as theoretical advantages over the mapping hypothesis (from now on referred to as MH). Making use of MH the difference in meaning would be associated with a difference at LF, rather than domain D.

2.1. Simple NPs

Let us start with simple NPs containing *einzig*:

(5) (a) *Peter ist der einzige Mann.*
'Peter is the only man'
(b) *Der einzige Lehrer ist gekommen.*
'The only teacher has come'
(c) *Maria hat den einzigen Pfarrer getroffen.*
'Mary has met the only priest'

In case of '*einzig*-NP' without any further (restrictive) modification we do not find any ambiguity. These NPs are interpreted w.r.t. a presupposed set, i.e. w.r.t. a set whose member share a common property that is somehow salient in the discourse. This is similar to what we found in case of stress on N. In this case it is also possible to have the set of alternatives not explicitly expressed, i.e. Identification is not necessary

syntactically. In case of *einzig* it is again remarkable that it is only the presupposed property but not the distinguishing property that need not be present syntactically, i.e. that can be salient in the discourse. Thus in case of simple NPs the property distinguishing one member of the set of individuals is necessarily the property denoted by N (namely being a man, a teacher or a priest). It is however necessary that the set of individuals we talk about has been established in the discourse or is at least recoverable from the context.

Within MH one could argue that this interpretation is induced through the presence of an operator. Instead of the determiner binding the variable provided by N it could be the operator which binds this variable. The fact that in case of a simple NP a previously established set is presupposed could be derived from an independent property of operator-variable constructions: Operators necessarily need a restrictor[2]. So we could assume the following interpretation:

(6) Op(einzig) x [x a member of set S] [x a man]

According to the analysis I am assuming the interpretation arises in a similar way as in case the NP is associated with focus. Since *einzig* is associated with the meaning given in (3) it forces the presupposition that there is a set of individuals out of which the particular individual is picked. Again the interpretation of X corresponds to the set of possible antecedents, i.e. the actual content *Mann* seems to be emptied. The denotation of N is interpreted as the distinguishing property for picking out the relevant discourser referent x. Thus we get a similar interpretation as in case of stress on N:

(7) *der MANN, der am Rennen teilnimmt*
the MAN, who at-the race participates

Here it is also the case that being a man is the distinguishing property, sufficing to pick out a particular individual which is the intended referent of DP. This was interpreted by claiming that there is an incomplete condition X = ? (i.e. an 'emptied' N). Moreover I argued that there is some denotation associated with this X, namely the *set of possible antecedents* relativized to the present discourse, as in case of pronominals.

[2] cf. Clark (1990) arguing for the following constraint:
The Restriction Constraint:
[YP Qx:...Z...] if and only if Z restricts the domain of quantification of Qx, where YP is an Operator.
Notice that Clark (1990) assumes that the restriction can be implicit.

2.2. Modified NPs

Consider now NPs containing a restrictive modifier:

(8) (a) *der einzige Mann, der am Rennen teilnimmt*
the only man, who participates at-the race
(b) *der einzige große Mann*
the only tall man
(c) *der einzige Mann mit einem Hut*
the only man with a hat

All these NPs are ambiguous w.r.t. the set of individuals we talk about. The NP in (8a) is ambiguous between talking about the *individuals participating the race* or about *a (certain) set of men*. The same is true for other restrictive modifiers. The NP in (8b) can either be interpreted so that we talk about a *set of tall people* one of them being distinguished through the property of *being a man*. On the other reading we talk about a (certain) *set of men* one of them being distinguished from the others by having the property of *being tall*. Consequently in (8c) it is either *the set of individuals wearing a hat* or *the set of men* we talk about.[3]

Following the analysis introduced above for simple NPs we derive this ambiguity through claiming that it is either NP (X) or DP (x) which is identified by the identifier (similar to what we have seen in section IV.7.). Alternatively according to MH we could say that the ambiguity arises by different mapping possibilities: it could ether be NP or the restrictive modifier respectively which maps into the nuclear scope. The optional mapping could be explained by claiming that the operator can bind the variable provided by the NP (resulting in the mapping of NP in the nuclear scope). Since the Operator needs a restriction and a restrictive element is present in the NP it is a natural assumption that this restrictive element maps into the restrictive clause. According to my analysis the option arises by the possibility to identify NP or DP (i.e. both X and x can be associated with an incomplete condition). Either property can be interpreted as being the distinguishing property (associated with the necessary uniqueness interpretation). Once the distinguishing property is established the other property necessarily will be interpreted as the property the individuals of the relevant

[3] Here we will again be mainly concerned with identifiers in the strict sense (i.e. relative clauses and 'complement-clauses'. Adjectives (as θ-identifiers) and PPs behave differently. The main difference is that they are not coindexed with D.

set have in common.[4]

2.3. Partitives

An interesting pattern arises if we look at *einzig*-NPs containing a partitive phrase:

(9) (a) **der einzige von den Männern*
the only of the men
(b) **der einzige der Männer*
the only the$_{GEN}$ men

(10) (a) *der einzige von den Männern, der am Rennen teilnimmt*
the only of the men who participates the race
(b) *der einzige der Männer, der am Rennen teilnimmt*
the only the$_{GEN}$ men who participates the race

This paradigm shows that *einzig*-NPs containing a partitive is ungrammatical in case there is no restrictive modifier. We have to explain the obligatory occurrence of an identifier. Notice that the present analysis provides an explanation for this behavior. It is an inherent property of the partitive phrase that it is necessarily interpreted as denoting the superset out of which the discourse referent is picked out. According to our analysis this means that it is necessarily associated with X, in fact it can be equated with X. So we can assume the following (abbreviated) representation:

(11) NP: X = set of individuals satisfying the property of being a man
PP: $x \subseteq X$
DP: x = ?

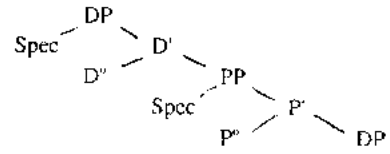
This representation captures the fact that the NP contained in the partitive phrase is (equated with) the superset. It plays the same role as X (i.e. the denotation of NP) in case of NPs without an overt partitive phrase. The difference to simple NPs is

[4] Note that here it is also possible to have a presupposed set of individuals in the restrictive clause (like in the case of the simple NP). In that case the relative clause can just be interpreted as non-restrictive. As we expect this reading becomes the only one available if the relative is forced to be interpreted as non-restrictive by adding *übrigens* ('by the way'):
(i) *der einzige Mann, der übrigens am Rennen teilnimmt*
the only man, who by the way participates the race
In this case the ambiguity is lost again, the NP behaves like a simple NP. Under our approach the non-restrictive relative simply disambiguates the NP: it can just be interpreted as denoting an individual. Thus there is no need for identification, since there is no incomplete condition.

however, that the subset relation is somehow overtly expressed by means of the PP. Therefore I assume that $x \subseteq X$ corresponds to the denotation of PP. Here it is explicitly expressed, in other cases it is just a byproduct of the fact that the discourse referent must be a member of the set denoted by NP, i.e. the subset relation is only implicit, rather than explicit. Given the representation above it is obvious why we need a restrictive modification. Similar to stress on the determiner *einzig* forces a uniqueness-interpretation. Since it is part of the meaning of the partitive NP that it has to denote the superset the property denoted there cannot serve as the distinguishing property (as in case of simple *einzig*-NPs). Notice moreover that the partitive phrase necessarily involves a plural NP (i.e. it must denote a set containing more than one member) another property has to be added in order to pick out one particular discourse referent unambiguously. This amounts to saying that x is not identified therefore we find the incomplete condition x = ?. This incomplete condition gives rise to the ill-formedness of *einzig*-NPs containing a partitive phrase but no identifier (i.e. no identifying property).

Notice that the assumption that the partitive phrase is equated with the denotation of the superset could imply the following (structural) property of such phrases. In case of simple NPs the denotation of the superset is NP, i.e. the complement of D. We might generalize this assumption. Assume that the superset for the denotation of D (x) is always associated with the complement of D. Let us assume that a partitive phrase is in fact the complement of D. This assumption leads to the following representation of DPs containing a partitive phrase:

(12)



Abstracting away from the position of *einzig*[5] we can assume that PP is a direct complement of D°. So we might put things the other way round. It might be a consequence of the syntactic position of the partitive (rather than an inherent property as suggested above) that it is necessarily interpreted as the superset for the denotation of DP (x). Notice that this analysis differs form previously suggested analyses in that I am not assuming an empty N between D and PP. Olsen (1991) for example suggests

[5] It can either be in a position adjoined to PP or alternatively it could be assumed that there is an AgrP between DP and PP (in the sense of Giusti 1991) that hosts *einzig*.

that there is an NP that is headed by an empty N. This empty N is assumed to be projected from the lexicon but lacking a phonological matrix. She assumes that the empty N takes the same denotation as the antecedent (which is in this case a 'postcedent'). However, there are problems with such an analysis that the representation assumed here does not face. First, under Olsen's analysis it would not be evident why *einzig*-NPs containing a partitive necessarily need an identifier. If there were an N (which lacks a phonological matrix) we would expect that this N could provide the superset for x. Under this assumption we would expect the same behavior of *einzig*-NPs containing a partitive and simple *einzig*-NPs. The only difference were that in the latter case the N denoting X has a phonological matrix. There is no reason why the possibility to in interpret X as being a presupposed set (of possible antecedents) should be restricted to overt serialization of N (to the contrary one might expect that this interpretation, which corresponds to an 'emptied' N, is not available if this N is (phonologically) empty, but this is apparently the case, as we have seen. If we assume that there is no NP involved we can explain the necessity for the restrictive modifier since any x must be associated with an X. If no N denoting X is present then the partitive phrase is interpreted as denoting X. Secondly, if there were an empty N (with all its features except for the phonological matrix) we would expect the same interpretation of an NP having this N overtly expressed:

(13) (a) *der einzige Mann von den Männern*
the only man of the men
(b) *der einzige Mann von den Männern, der einen Hut trägt*
the only man of the men who wears a hat

It is true that the same reading as in examples (9-10) above is available (in this case a restrictive element is again necessary). However the NP in (13a) is not illformed (as opposed to (9)). However, it is only well-formed under a particular reading that is not available in case N is not overt. This is a reading where the denoted man is interpreted as being *the only real man* of the set of men denoted in the partitive phrase. Although all the individuals have to property of being a man there is actually just one of them that has the property of being a real man. This interpretation should become clear on basis of the following example:

(14) *Peter ist der einzige wirkliche Mann von diesen Männern, (die ja alle nur Schlappschwänze sind).*
Peter is the only real man of these men (who are all just softies)

This difference in interpretation depending on the overtness of N is not expected under an analysis à la Olsen. Under the present analysis this behavior follows. In case there is no overt N the partitive phrase is assumed to be a direct complement of D. If N is overtly expressed, then it is in the complement position. However in this case it is interpreted as the distinguishing property. In this case the superset (i.e. the set of alternatives) can be either salient in the discourse or overtly expressed. The partitive phrase is one possibility to doing so (i.e. it can identify X).

2.4. Abstract Nouns

So far we have seen the behavior of *einzig* w.r.t. *concrete* individuals. If we look at abstract nouns we find a somewhat different behavior:

(15) (a) *das einzige Argument*
the only argument
(b) *die einzige Hoffnung*
the only hope

If we compare the interpretation of these *einzig*-NPs with the interpretation of other *einzig*-NPs we find an interesting distinction. Consider again the interpretation associated with *einzig* repeated here for convenience:

(16) There is a (presupposed) set of individuals and one member of these individuals is distinguished from the others by being the only one who fulfills the property P.

In case of abstract nouns we cannot really talk about a presupposed set of individuals one of them being distinguished through *being an argument*. What else could the interpretation be in these cases? First notice that it has to be the case that the denoted discourse referent must be a member of a presupposed set of individuals. This can be shown on basis of the ill-formedness of *einzig* modifying Names:

(17) (a) **der einzige Peter*
the only Peter
(b) **das einzige Paris*
the only Paris

So if the reference of the individual/entity is inherently specified, i.e. if there is no need for Identification then one cannot get the interpretation that you refer to a single individual which is distinguished from others by having the property N. Therefore this phrase results in ungrammaticality.

Now we have to solve the problem that on the one hand the behavior of Names shows that we have to assume that a presupposed set is available and on the other hand abstract nouns seem to resist an interpretation where a set of given entities is presupposed one member of this set being distinguished by the property denoted by N. The problem is that we have to find a denotation for X in case N is the distinguishing property. Intuitively the interpretation of an NP containing an abstract N and *einzig* can be paraphrased as follows:

(18) Of the set of possible Ns there is just one which has the property of being a real (or the only existing) N.

It seems to be the case that the individual denoted by an abstract *einzig*-NP is contrasted to a set of individuals containing possible (or existing) N's. To make this intuition more clear consider the NP in (15a) *the only argument*. It can be used in two different contexts. The first context is as follows. Assume that there is a discussion about the possibility that Peter is a spy. Everybody is arguing for or against it. All the arguments presented are rejected on basis of some reason or other except for one particular argument. This is then the only remaining argument and therefore it is the only real argument of the set of possible arguments. As we have seen in examples (13-14) the interpretation of being the only *real N* is also available with concrete nouns.

[5] Note that this example is well-formed if the name is interpreted as a property. The two possible interpretations of names can be observed in the following paradigm.
(i) *(der) Peter ist gekommen*
(the) Peter is come
(ii) *Dieser Mann heißt (*der) Peter*
This man is named (*the) Peter
The possibility for a definite determiner shows that names can also be interpreted as properties (namely the property of being named 'Peter'). This is exactly the interpretation if *einzig* modifies the name. In that case the presupposition arises that there are other individuals that have the property of *being named Peter*.

An NP like *the only man* can have the interpretation that there is a set of men and just one of these men has the property of being a real man (meaning that this is the only one who is totally male).

The other possible interpretation for *the only argument* is slightly different. It is in principle possible that there are several arguments for a given fact but in the actual world there happens to be just one argument. This amounts to saying that out of the set of possible arguments there is just one (existing) argument. The two interpretations have something in common: namely that the entity referred to is contrasted to a set of individuals that can either exist or their existence might be possible. However all of them have the same property in common, namely *being an argument*. This means that in case of abstract nouns we find a similar interpretation as in case of pronominals. Remember that I have argued that pronominals are associated with an NP that denotes X. X corresponds to the set of possible antecedents. Here we get a similar interpretation. X corresponds to a set of possible discourse referents and this set is furthermore restricted by the denotation of N: it must be the set of possible arguments.

2.5. Possessor phrases

I will now look at restrictions governing Identification, i.e. syntactic environments where the ambiguity detected so far is lost. The first environment we will look at is possessor constructions. Here we find a contrast w.r.t. pre- vs. postnominal possessor phrases. Consider the following paradigm:

(19) (a) *das einzige Bild Peters*
the only picture Peter's
(b) *das einzige Bild von Peter*
the only picture of Peter

The NPs in (19) with the possessor in postnominal position (realized as a GENNP (19a) or as a PP (19b)) are ambiguous between the following two readings:

(20) (a) of the set of entities possessed by Peter, x is the only one that has the property of being a picture
(b) of the set of pictures (we talk about), x is the only one that belongs to Peter

According to what we have said so far this ambiguity arises as follows: it is either the property denoted by N that is interpreted as the distinguishing property (and therefore the identifying property) or it is the property of *being possessed by Peter* that is distinguishing, and therefore identifying x out of the set X. Interestingly, in case of prenominal possessor phrases this ambiguity is lost:

(21) (a) *Peters einziges Bild*
Peter's only picture
(b) *dem Peter sein einziges Bild*
the$_{DAT}$ Peter his only picture

Here the only reading available is the one where we talk about *individuals possessed by Peter*. In this case the property denoted by N must be the distinguishing property, i.e. the denoted discourse referent is distinguished by *being a picture*. The other reading (which is available with postnominal possessor phrases) where we would talk about a (certain) *set of pictures* is not available. Why could that be? The analysis of Identification suggested here provides an explanation. The distinguishing property can be equated with being the identifier. The identifyee has been equated with x and x is supposed to be the denotation of DP. Prenominal possessor phrases thus receive the following representation:

(22) $[_{DP}$ Peter $[_{D'}$ $[_{Do}$'s] einziges $[_{NP}$ $[_{No}$Bild]]
D: x = ?
N: X = set of pictures

Remember that Identification is subject to PCI: The identifier can not precede the identifyee. Since it is D that is associated with x (the index of the identifyee) the possessor in (22) linearly precedes the identifyee. Therefore it cannot be interpreted as the identifier. Being the identifier is however equated with denoting the distinguishing property, i.e. the property that should suffice to pick out the relevant denotation for x. It follows that in case of prenominal possessor phrases the ambiguity is lost: Being possessed by Peter cannot be interpreted as the identifier for x and therefore it cannot be interpreted as the distinguishing property.

If this is indeed the explanation for the loss of ambiguity we cannot analyze the existing reading by claiming that the possessor identifies NP (X). I do not think that this is a problem for the present analysis. I have argued that Identification provides a licensing configuration. In case of (22) the possessor is inherently licensed by means of the Specifier Head relation, which counts as a licensing configuration. There is no

need to establish a relation of Identification in order to license the Possessor. However if the prenominal possessor is not licensed as an identifier (neither of X nor of x) then the loss of ambiguity cannot be explained by saying that it can only identify one of the two possible identifyees (X or x). It must be related to some other factor.[7] I think that this is indeed the case. *einzig* necessarily induces the uniqueness interpretation. Therefore there must be some property that distinguishes x out of X. If the possessor is in prenominal position it cannot count as the distinguishing property. The only property available is the one denoted by N, similar to simple NPs discussed in section 2.1. There we have seen that the set out of which x is taken can be presupposed in the discourse. If there is a possessor phrase the most natural interpretation for this set is that it contains entities possessed by Peter since x is possessed by Peter.

If the possessor occurs in postnominal position, then it is possible to interpret it as an identifier of either X or x. In this case we expect LCI to constrain its position. In case the identifying possessor phrase occurs in extraposed position we expect that Identification of X is no longer available (DP intervenes). This means that the only available interpretation is that the property denoted by N is the distinguishing property:

(23) (a) *Ich habe das einzige Bild von Peter gekauft*
I have bought the only picture of Peter's
(b) *Ich habe das einzige Bild (gestern) gekauft von Peter.*
I have bought the only picture (yesterday) of Peter's

The prediction is borne out: the ambiguity is lost if PP is extraposed. This follows from LCI because the identifier is too high in order to have access to NP, it can just identify DP.

The same effect is also found with other instances of Identification. If we are dealing with an NP containing *einzig* construed with a relative clause in extraposed position, then we expect no ambiguity. This prediction is again borne out:

[7] Notice that a similar phenomenon is also found in case of other NPs (not containing *einzig*). Consider the following minimal pair, involving either a pre- or a postnominal possessor and an adjective:
(i) *das große Buch Peters*
(ii) *Peters großes Buch*
In (i) the adjective can either be restrictive or non-restrictive, whereas in (ii) it can only be interpreted as non-restrictive. I do not have anything to say about this fact.

(24) (a) *Der einzige Mann der an dem Rennen teilnimmt, ist verschwunden.*
The only man who participates the race has disappeared
(b) *Der einzige Mann ist verschwunden, der an dem Rennen teilnimmt.*
The only man has disappeared, who participates the race

In (24a) the interpretation of the DP gives rise to two different readings: either *being a man* or *being a participant in the race* can be the distinguishing property. In sentence (24b) the ambiguity is lost. The only available interpretation for the DP is that *being a man* is the distinguishing property. The loss of ambiguity follows from the present analysis. In extraposed position the possibility to identify NP (X) is no longer available. LCI would be violated. X is not accessible for Identification.

Notice that the two cases where we find a loss of ambiguity (i.e. prenominal and extraposed possessor phrases) crucially support the claim that the two readings are relevant for Grammar. The two different readings should be encoded. Syntactic configurations have an influence on the availability of the two readings. Moreover the relevant interpretation of the DP is sensitive to syntactic configuration in a way that is predicted by the present analysis of Identification: both, if PCI and LCI are violated respectively then one of the two readings is lost.

3. Superlatives

A similar behavior is found w.r.t. superlatives. (Notice that superlatives are also carrying a uniqueness presupposition - as we have already mentioned in section IV.3.) Consider the following DP:

(25) *der schönste Mann*
the most beautiful man

This DP is ambiguous in a similar way as *einzig*-NPs involving restrictive modification are. The ambiguity is as follows. Under the first reading a certain *set of men* is presupposed. One of these men is distinguished from the others by *being the most beautiful of them*. Under the second reading (which is more easily available with (contrastive) stress on N) we talk about the set of *most beautiful individuals*. One individual out of this ist is distinguished by the property of *being a man*. The question is now how this ambiguity arises and why in this case (as opposed to *einzig*-NPs) there is no need for restrictive modification in order to get two readings. The difference is that the adjective itself (i.e. the element that induces the uniqueness interpretation)

provides a property that can serve as the distinguishing property. *einzig* does not provide such a property. Interestingly, there is a (derived) adjective denoting the property of being *einzig* (i.e. unique):

(26) *der einzigartige Mann*
the outstanding man

4. Indefinite NPs

So far we have not said anything about indefinite NPs and Identification. Notice that identified indefinite NPs do not create a problem for the present analysis. They can introduce a discourse referent (therefore obeying the Novelty Condition resulting in the precedence constraint). It can be assumed that the identifier provides availability in the discourse by ascribing a certain property to the denoted discourse referent. An interesting pattern arises if we look at indefinite *einzig*-NPs. Consider the following sentence:

(27) *Maria hat [einen einzigen Mann, der am Rennen teilnahm] gesehen.*
Mary has an only man seen, who at the race participated seen
'Mary has seen a single man who participated the race'[8]

The DP in (27) still refers to a single individual like in case of definite NPs. However, there is still a difference to definite DPs w.r.t. what the distinguishing property can be. This behavior sheds some light on the behavior of indefinite NPs in general. The difference is as follows. An indefinite *einzig*-NP' gives rise to the following presuppositions:

(28) (a) There are other men seen by Mary that did not participate the race.
(b) There are other individuals seen by Mary, that are not men.
(c) There are other tall men, that were not seen by Mary.

The presuppositions in (28a) and (28b) correspond to the ones we have seen in case of definite *einzig*-NPs. The crucial difference between indefinite and definite DPs is the presupposition in (28c) which is not available for definite *einzig*-NPs. The distinguishing property can not only be the properties found within the DP itself. The DP seems to be transparent w.r.t. the matrix clause, i.e. the property expressed by the

[8] Note that English and German differ in that English does not allow *only* (corresponding to German *einzig* otherwise) with indefinite NPs.

whole proposition has influence on the possible presuppositions. The denoted individual (*the man participating the race*) has as its distinguishing property that *he was seen by Mary*. It can be the case that there are other *men who participated the race* but there is only one of these men who was seen by Mary. The individual denoted by an indefinite NP is the only one who fulfills all three (available) properties (being a man, participating the race and being seen by Mary).

In case of definite NPs the third reading is not available: the denotation of the DP is independent of any other property available in the sentence. In *Maria hat den einzigen Mann, der am Rennen teilnahm gesehen* ('Mary has seen the only man who participated the race') one can not presuppose that there are more individuals who fulfill the properties of being a man and participating the race. It is not possible to interpret the property denoted in the whole proposition as the distinguishing property. This amounts to saying that indefinites are transparent whereas definites are not. This is however not a surprising result. Notice that the transparency or opacity of indefinites vs. definites respectively can be observed for various phenomena (for example movement is in many cases possible out of the former but not out of the latter).

Although I do not want to suggest any further solution within the present analysis I still think that this transparency found here sheds some interesting light on the difference between 'strong' and 'weak' NPs. Assume that the former do induce a particular discourse referent, which can be fixed just on basis of the DP. The latter however cannot be fixed only on basis of the DP itself but only in combination with the sentence it occurs in. This would then correspond to 'nonspecific' readings. I will not pursue this issue any further since it goes beyond the scope of the present discussion. However, one more issue needs to be addressed. As we have seen there is a way to translate the present analysis into the MH. This would then mean that there is a correlate to the MH DP-internally. Moreover we have seen that the different interpretations have to do with presuppositions - and as a matter of fact Diesing (1992) talks about a *presupposed reading* for specific NPs. However, we have also seen that in case of DPs there is no one to one correlation between the position of an element and its interpretation. Elements that occur in the restrictive clause do not show any restriction on their position (in fact they need not even be expressed overtly), whereas elements to be mapped into the nuclear scope DO show such a restriction. This would suggest that *specific* NPs in the sense of Diesing should be allowed in any position whereas *non-specific* NPs should be more restricted. I will leave the problem of a DP-internal MH and its implications for IP-MH à la Diesing (1992) as a matter of future research.

VII. A PROBLEM FOR PCI: LEFT DISLOCATION

The analysis discussed so far has to face one particular problem, which I will address in this chapter in more detail. It will turn out that once an appropriate analysis is established the problem disappears. Moreover we will see that the data under consideration will turn out to confirm our analysis. The particular problem has to do with the Precedence Constraint (derived by the Novelty Condition) assumed for Identification:

(1) The Precedence Constraint on Identification (PCI)
The identifyee has to precede the identifier.
(where XP precedes YP iff it linearly precedes every member of YP).

Consider again the paradigm derived by PCI:

(2) (a) *Peter hat [[daran]$_i$ [daß Maria Bier trinkt]$_i$] geglaubt.*
Peter has thereon, that Mary beer drinks believed
(b) *Peter hat [daran]$_i$ geglaubt, [daß Maria Bier trinkt]$_i$*
Peter has thereon believed, that Mary beer drinks
(c) **[Daß Maria Bier trinkt]$_i$ hat Peter [daran]$_i$ geglaubt.*
That Mary beer drinks has Peter thereon believed
(d) *[Daß Maria Bier trinkt] hat Peter geglaubt.*
That Mary beer drinks has Peter believed
(e) *[[Daran]$_i$ [daß Maria Bier trinkt]$_i$] hat Peter geglaubt*
Thereon that Mary beer drinks has Peter believed

The identifier can occur in a position right-adjoined to the identifyee (2a) and (2e) and in extraposed position (2b). But it cannot be moved to a position where it precedes the identifyee (2c) although it is in general possible to move the potential identifier to this position if no identifyee is present (2d).

Consider now the following sentence contradicting the claim that the identifier cannot be in a position preceding the identifyee:

(3) *[Daß Maria Bier trinkt]$_i$ [daran]$_i$ hat Peter geglaubt.*
that Mary beer drinks thereon has Peter believed

In the sentence above the identifier is in a position preceding the identifyee (i.e. the pronominal element). The only difference between (3) and the ungrammatical (2c) is

that in (3) the identifyee occurs in topicalized position. This configuration is known as *Left Dislocation*. The wellformed example (3) shows exactly the same pattern as other instances of Left Dislocation involving nominal elements (instead of a left dislocated sentential constituent):

(4) *[Den Peter]$_i$ [den]$_i$ habe ich nicht gesehen.*
The Peter d-pron have I not seen

In this section I will argue that the apparent problem can be dispensed with by claiming that in case of Left Dislocation we are not dealing with an instance of Identification. Thus the main claim will be that the left dislocated constituent acts as a discourse antecedent for the pronominal element, i.e. there is no direct dependency relation between the left dislocated element and the topicalized pronoun. Remember that I have argued in section III.5.3. that discourse antecedence and syntactic antecedence are independent of each other. Both notions of antecedence should not be collapsed. With respect to anaphoricity I would like to argue that the sentence above is similar to intersentential anaphoric relations like in the discourse below:

(5) *Klara hat gestern [den Peter]$_i$ gesehen. Den$_i$ hab ich nicht gesehen.*
Klara has yesterday the Peter seen. D-pron have I not seen.
'Clair has seen Peter yesterday. I didn't see him.'

Here the occurrence of the DP *den Peter* in the first sentence is as a discourse antecedent for the pronominal element in the second sentence. There is no direct syntactic dependency relation involved. Notice moreover that Left Dislocation necessarily involves a pronoun that is rather similar to a demonstrative pronoun. This was also argued to be a property of identifyees (which just need Identification if the content is not clear in the context). We will see that Left Dislocation is somehow the *opposite* of Identification (not only because Left Dislocation involves a leftward relation whereas Identification involves a rightward relation). This will be confirmed by the presupposition induced by Left Dislocation. The presuppositions arising in case of Left Dislocation will be argued to be the exact opposite to the ones we have seen so far (being induced by Identification). Both presuppositions have to do with the index X associated with the pronominal. In both cases the pronominal makes the denotation associated with X (namely the set of possible antecedents) available.

In the following sections I will present the analysis for Left Dislocation I am

assuming[1]. We will see that assuming no direct (syntactic) dependency relation (which I am forced to assume to get rid of the apparent problem) is independently motivated (both theoretically and empirically). We will then be able to compare extraposition and Left Dislocation. Many properties of Left Dislocation will confirm the assumptions we have discussed so far. We will look at the different presuppositions and how they follow from the present analysis. We will also see some facts supporting the Optionality Hypothesis (as introduced in section III.3.). Finally Left Dislocation will provide evidence for the claim that properties are associated with discourse referents.

1. Some problems concerning Left Dislocation

Let me start with looking at an instance of Left Dislocation in more detail. Consider the following example:

(6) *[Den Frosch]$_i$, [den]$_i$ hat die Prinzessin nicht geküßt.*
The frog, pron. has the princess not kissed
'The frog, the princess didn't kiss him'

What we find here is a DP (*den Frosch*) which is left-adjoined to a fully saturated V2 clause (*den habe ich nicht gesehen*). In this clause, a pronoun occurs in topicalized position i.e. SpecCP, and this pronoun is co-referential with the left-dislocated DP. I will assume without further discussion that the left-dislocated constituent is *basegenerated* in a position *adjoined to CP*. Therefore we are in the need to address the question as to how it gets licensed in this position. Another problem to be addressed is the requirement that the left-dislocated constituent and the topicalized pronoun match in certain features. We will see that there is a single answer to both questions while we can still assume that the left-dislocated constituent acts as a discourse antecedent for the pronoun.[2] There is one more property of Left Dislocation we have to draw our attention to. The sentence in (6) is perfectly grammatical even without the left-adjoined DP as shown in (7):

[1] cf. Wiltschko (1994) for a more detailed discussion of this analysis.

[2] Notice that anaphoric resolution in general is just possible if the two elements match in certain features (i.e. number and gender). So this is not a peculiarity of Left Dislocation but also of intersentential anaphoric resolution. However, the left-dislocated constituent and the topicalized pronoun match in more features than intersentential anaphorically related constituents do: both elements have to bear the same Case. Moreover, in case of a left-dislocated PP the preposition has to also be found in the topicalized constituent.

(7) *Den hat die Prinzessin nicht geküßt*
pron. has the princess not kissed
'The princess didn't kiss him'

Given the optional occurrence of the left-dislocated element we have to look for a possible licensing relation that is not obligatorily established but one which can be activated in case it is needed. In other words the CP to which the left dislocated constituent is adjoined has to provide an environment where the left dislocated constituent can be licensed.

2. The Proposal

The analysis for Left Dislocation I will assume (following Wiltschko 1994) can be summarized as follows. The left dislocated constituent is licensed as a syntactic operator (a notion which will be clarified shortly) which is *parasitic* on the (syntactic) operator-variable-chain between the topicalized pronoun and its trace. The (simplified) representation for Left Dislocation is given in (8) (where POp refers to *Parasitic Operator* and TOp refers to *Topic Operator*):

(8) DP_i $[_{CP}$ $Pron_i$ ..... t_i
POp TOp vbl

It is crucial that there are two seperate chains involved both of them containing the same trace: $C_1 = \{TOp,t\}$, $C_2 = \{POp,t\}$. The two chains differ in that C_1 is derived by movement (i.e. topicalization) whereas C_2 is basegenerated. POp is assumed to be parasitic on C_1. Notice that this analysis captures the traditional problem that Left Dislocation shows properties of basegeneration AND of movement. Given this representation there are several questions that arise immediately: What exactly do we mean by Parasitic Operator? Why is there no chain between POp and TOp? Assuming no chain between POp and TOp seems to indicate a problem for an account of the feature-sharing between POp and TOp. We will also address this question.

It should be clear now that this analysis provides a solution for the apparent problem concerning Identification and PCI. PCI is not violated since there is no dependency relation between the two elements under consideration. The left dislocated constituent behaves exactly like a discourse antecedent. We have seen in discussing

Identification, that there need not be a syntactic identifier. If the content of the pronominal is clear from the context, i.e. if it is preceded by a discourse antecedent (that is still available). Since we are not dealing with an instance of Identification LCI is also not relevant in this configuration. The locality constraint is captured differently: it follows from the fact that the left-dislocated element is parasitic on another A'-chain. The only possibility to create such a chain is by means of topicalization of the coreferent element. Topicalization in German targets SpecCP which then happens to be adjacent to the position of the left-dislocated element.

3. Motivating the analysis

In this section I will show that the analysis for Left Dislocation assumed here does not only solve the problem concerning PCI but it is independently motivated both, theoretically and empirically.

3.1. Theoretical motivation

Being licensed as a parasitic operator is the only possibility for the left dislocated element to be licensed. Several theoretical considerations justify this claim. Moreover I think that the assumptions concerning the general theory of licensing discussed in chapter I. are relevant here as well. They are not designed just to explain the problems concerning extraposition.

3.1.1. Being in an operator-variable chain is the only possibility for the left dislocated constituent to be licensed

As we have already discussed in chapter I. we can distinguish between two different kinds of licensing requirements: There is inherent licensing, i.e. licensing just by virtue of sitting in a certain (licensed) position and there is relational licensing, i.e. elements which are licensed by another element. Relational licensing is mostly constrained by certain locality requirements governed by c-command. Remember moreover that I am assuming that all of the relational licensing requirements are asymmetric relations, i.e. we never find mutual licensing relations. Given that a left dislocated constituent cannot be licensed relationally since there is no element asymmetrically c-commanding it we already have a reason to assume that left dislocated constituents are licensed inherently.

Remember that I have argued in chapter I. that inherent licensing can be equated with Chomsky's (1992) notion of legitimate-LF-objects, i.e. heads, arguments, modifiers and operator-variable-chains. It is neither the case that a left dislocated constituent is a head nor an argument, nor a modifier - so we are left with the last possibility: what we find is an operator-variable chain. Therefore we can say that the theory of licensing in general forces us to arrive at the conclusion that left dislocated constituents are licensed as operators, which is exactly the analysis suggested here.[3]

3.1.2. What is a Parasitic Operator?

In the literature we find two different definitions for what counts as an operator. One possibility is to give a list of elements functioning as operators (cf. Chomsky (1981), Cinque (1990)):

(9) Operator $=_{def}$ bare quantifiers, wh-phrases, null NPs in SpecCP

The other possibility is to define operators structurally:

(9') Operator $=_{def}$ the head of an A'-chain

I will adopt the intensional (structural) definition in (9'). Still one has to account for the distinguishing property of the elements listed in (9). These elements are the ones which are forced to move to an A'-position - presumably because of the prohibition against vacuous quantification, i.e. they have to bind a variable). W.r.t. heads of A'-chains in general (which are not elements of the set defined in (9)) it seems that there is nothing inherent about these elements that forces A'-movement. However, as soon as a constituent moves to an A'-position it acts like an operator by virtue of binding a variable (left behind through A'-movement). Therefore one can distinguish two kinds of operators: *inherent operators* and *syntactic (derived) operators*. In other words: inherent operators are only a subset of operators.

It has been argued in several places that one has to make use of this distinction (cf. Sportiche (1983), Lasnik & Stowell (1991)). Moreover it has been argued there

[3] cf. among others Cinque (1990), Dimitriadis (1994), Anagnostopoulou (1994), Iatridou (forthcoming) for a discussion of the operatorhood of left dislocated and clitic left dislocated constituents.

that the *bijection principle*[4] only holds for inherent operators but not for syntactic operators. It is claimed that a syntactic Operator can bind more than one variable. However, if the bijection principle can be violated in one direction there is no reason why it should not also be possible to violate it in the other direction: one variable being bound by two operators. This is what I am suggesting here. The left dislocated constituent is licensed by virtue of being linked to a variable, i.e. an A-position.[5] This linking is however only possible if there is an adequate variable available in the clause the left dislocated constituent is adjoined to. Notice that the cross-linguistic difference between CLLD and Left Dislocation can be assumed to follow from a lexical property[6], i.e. clitics can act as bound variables therefore they can stay in situ and thereby licensing the CLLDed constituent. d-pronouns on the other hand cannot act as bound variables themselves. Notice that this is an independently attested property of d-pronouns as shown in the following example:

(10) *[Jeder Prinz]$_i$ glaubt daß [der]$_{*i/j}$ [er]$_{i/j}$ eine Prinzessin kriegen wird.*
Every prince believes that d-pron/he a princess get will

The d-pronoun cannot be coreferent with a quantified antecedent, i.e. it cannot receive a bound variable interpretation. Given this independently attested property of d-pronouns it follows that in order to provide an licensing configuration for left dislocated elements they have to A'-move in order to leave a variable. There is still another question arising now: why is it not possible to have a left dislocated constituent licensed by means of binding a (personal) pronoun (that can act as a bound variable) in its base-position. Notice that such a configuration is indeed possible, however, it gives rise to a *Hanging Topic* interpretation. I will briefly address this question in section VII.5.

[4] The Bijection Principle basically says that there is a 1:1 mapping between Operators and variables: an Operator can bind only one variable, and a variable is bound by only one Operator.

[5] Assuming Sportiche's (1983) and Lasnik & Stowell's (1991) conclusion that the bijection principle holds only for inherent operators as well as the assumption made here that both the left dislocated constituent and the topicalized pronoun are syntactic operators this analysis derives the fact that Left Dislocation is just possible for elements which are not inherent operators. Left Dislocation is not possible for elements carrying quantificational force:

(i) **Keinen Frosch, den hat sie geküßt*
No frog, pron. has she kissed

(ii) **Nicht den Frosch, den hat sie geküßt*
Not the frog, pron. has she kissed

(iii) **Niemanden, den hat die Prinzessin geküßt.*
Nobody, pron. has the princess kissed

[6] cf. Anagnostopoulou (1994) for such an analysis.

3.1.3. Why is there no chain between POp and TOp?

I will now show why we are theoretically forced to assume that there is no chain between POp and TOp and therefore it can be assumed that POp acts as a discourse antecedent for TOp. There is no direct dependency relation between the two elements under consideration. Consider again the abstract representation of Left Dislocation suggested here:

(11) 

Recall that we make use of a structural (as opposed to an extensional) definition for operators (being the head of an A'-chain). If we would assume that there is just one single chain, then just POp would be an operator in the representation in (11). There is no way in which TOp could be an operator since it is not the head of an A'-chain rather it would itself be bound by POp. However, if TOp is not an operator but rather an *intermediate* binder the variable would not be locally A'-bound by an operator. TOp would be a closer A'-binder without being an operator.

Notice that the same kind of technical problem arises with intermediate traces. Consider the example below (taken from Cinque 1990):

(12) *Who$_i$ did he say [t'$_i$ [t$_i$ was invited]]*

Here the variable is not locally A'-bound by an operator since t'$_i$ is the local A'-binder. There are two different proposals to be found in the literature in order to deal with this problem. Cinque (1990) argues on basis of this example for the following definition of variables:

(13) Variable $=_{def}$ [$_{NP}$e] in A-position locally A'-bound and operator bound.

The other possibility is to assume that intermediate traces are deleted at LF (cf. Lasnik & Saito 1984 among others). Assuming this kind of deletion no problem arises with respect to binding at LF.

However, neither solution is available in case of Left Dislocation. First reformulating the definition of variables (along the lines of Cinque 1990) does not help

to allow the topicalized element to count as an operator. Secondly, within a deletion account one would have to claim that the topicalized pronoun deletes at LF yielding the following LF-representation:

(14) *Den Frosch habe ich geküßt.*
the frog have I kissed
POp$_i$ t$_i$

Here no problem w.r.t. binding would arise. Notice that this representation looks like the output of topicalization. In the following sections I will present arguments against such an analysis.

3.1.4. The pronoun cannot be deleted

There are several arguments against deletion of the pronoun. First of all it is not true that topicalization and Left Dislocation have exactly the same interpretation (but they should have if they had the same LF representation). We will see the differences in interpretation in section VII.5., they have to do with different presuppositions depending on the presence or absence of the pronominal respectively.[7] Moreover, in some cases the pronoun in topicalized position is even necessary to yield a well-formed output:

(15) *Daß der Frosch ein Prinz ist *(dessen) war sich die Prinzessin nicht bewußt.*
that the frog is a prince (ProPP) was herself the princess not aware
'The princess was not aware of the fact that the frog was a prince.'

The ungrammaticality of (15) without the d-word is probably due to selectional requirements of the verb. *bewußt sein* ('to be aware of') selects for a genitive NP. So if we had to assume that the d-word deletes at LF in order to get the right representation we would have to delete a selected argument. This would yield a violation of the Projection Principle as well as a violation of the condition on the recoverability of deletion.

[7] To be exact, given the present assumptions concerning the distinction between domain D and LF we might say (at least in the instances that only have an effect on presuppositions) that deletion at LF is an option but the pronoun has to be present at domain D (the level where I assume presuppositions to be encoded). However, in some of the following examples it will become clear that deletion at LF is not an option since it would violate the constraint on *recoverability of deletion* as well as the *Principle of Full Interpretation*.

3.1.5. Other instances of Left Dislocation

There is another crucial argument against deletion of the topicalized element which is at the same time an argument for the analysis presented here. It is a crucial aspect of my analysis that Left Dislocation is parasitic on topicalization. So far we have seen instances of topicalization of a coreferent pronoun. However, if it is true that Left Dislocation is parasitic on topicalization in general we expect that this does not just hold for topicalization of a pronoun. Left Dislocation should rather be possible in any environment of topicalization. The only constraint holding for the relation between the left dislocated constituent and the coreferent element is that they bear the same index. This follows from my analysis since both elements are assumed to be licensed by being mapped to the same position. Thus both constituents have to carry the same index as the variable occupying this position. Therefore we predict that any two constituents bearing the same index (i.e. denoting the same discourse referent) can occur in Left Dislocation environments.[8] This prediction is indeed borne out. Moreover, it shows that it is not just in case of Identification of a pronominal when Left Dislocation contradicts PCI. All the cases where I argued that we find an instance of Identification are possible in this 'construction'. I think that this confirms the claim that Left Dislocation and Identification are two (opposite) sides of one coin.

3.1.5.1. Two NPs

First consider two NPs denoting the same individual:

(16) (a) *[Die Prinzessin]$_i$ [dieses hübsche Mädchen]$_i$ hat den Frosch geküßt.*
the princess, this pretty girl has the frog kissed
'The princess, this pretty girl has kissed the frog.'
(b) *[Den Prinzen]$_i$ [den armen Schlucker]$_i$ hat die Prinzessin geküßt.*
the prince-acc, the poor guy-acc has the princess kissed
'The prince, the princess has kissed the poor guy.'

Left Dislocation is possible not only if a pronoun is topicalized but also in case a coreferent lexical NP is topicalized. Again topicalization is a necessary prerequisite for the left dislocated element to be licensed. Notice also that a lexical NP can in general

[8] As we will discuss in the next section this statement is too general but the restrictions on what elements can occur instead of the d-pronouns will be shown to follow from more general principles.

not act as a bound variable. Therefore it is again expected that the coreferent NP has to move, like in case of d-pronouns.

The examples above give rise to an interesting issue which also confirms the claim that the left dislocated constituent really acts as a discourse referent for the topicalized constituent. As Wasow (1972) observes, there is a *Novelty Condition*[9] for anaphoric relations:

(17) *An anaphor may not introduce any presupposition not associated with its antecedent.*

This condition captures the fact that an anaphorically dependent element cannot be more determinate in reference than its antecedent. It is crucial however that this is a nonstructural property - it also applies intersententially. Williams (1994) gives the following example:

(18) (a) *A captain$_i$ walked into the room. The officer$_i$ at first said nothing.*
(b) **An officer$_i$ walked into the room. The captain$_i$ at first said nothing.*

The NP *a captain* determines the referent more specifically than *the officer*. Wasow's Novelty Condition is violated in (18b) because the latter precedes the former. The same generalization holds for the above mentioned instances of Left Dislocation:

(19) (a) **Das Mädchen$_i$, diese Prinzessin hat den Frosch geküßt.*[10]
the girl, this princess has the frog kissed
(b) **Den armen Schlucker$_i$, diesen Prinzen hat die Prinzessin geküßt*
the poor guy, this prince has the princess kissed

The set of princesses is a subset of girls and therefore determines the referent more specifically violating the Novelty Condition.

Given these facts I think that we have further evidence that the element in SpecCP cannot be deleted. We have another piece of evidence that we are dealing with

[9] One should not be confused by the term 'Novelty Condition' as used by Wasow. It is different to Heim's (1982) notion of the Novelty condition which we were using so far. However both approaches are combined in the present analysis by means of claiming that Identification licenses (re-)introduction of a new discourse referent.

[10] Notice that the examples in (19) cannot be interpreted as instances of *parenthetical* rather than Left Dislocation. (see Wiltschko 1994 for a discussion)

two independent chains. It has to be noted that the examples discussed in this section recall Hirschbühler's (1975) arguments against analyzing Left Dislocation as an instance of a *copying transformation*. He crucially relies on the possibility for *epithets* to occur in the position where usually a pronominal element occurs. The following examples are taken from Hirschbühler (1975):

(20) (a) *Paul, Pierre s'est battu avec lui*
'Paul, Peter had a fight with him'
(b) *Paul, Pierre vient de se battre avec cet idiot.*[11]
'Paul, Peter has just fought with that idiot.'

Hirschbühler (1975) also mentions a constraint on what kind of NP can serve as an anaphor for other NPs which is reminiscent of our examples in (18-19). In fact it is quite similar to Wasow's Novelty Condition. He refers to Lakoff (1968) who shows that:

> *"there exists a hierarchy among different kinds of NPs, so that an NP_i may serve as an anaphor for an NP_j only if NP_j is lower in the hierarchy than NP_i".*
> (Hirschbühler 1975: 160).

Hirschbühler gives the following example as evidence for this constraint:

(21) **Cette tête-là, je pense que j'ai déjà vu cette grande blonde-là quelque part.*
'That face, I think I've already seen that tall blonde somewhere'.

We can analyze the NP that occurs instead of the d-pronoun in case of Left Dislocation as an epithet, i.e. the German counterpart of the French data discovered in Hirschbühler (1975).[12] Given that epithets (and according to Lasnik & Stowell (1991) almost any simple definite NP can) function quasi-pronominally (having a discourse antecedent) the facts discussed in this section are predicted by our analysis, which heavily relies on the fact that the possibility for Left Dislocation is parasitic on topicalization in

[11] Notice that in Hirschbühler's examples all epithets are DPs with a demonstrative determiner. Therefore the fact that demonstrative determiners for the topicalized NPs are also better in German seems to be a phenomenon which is not just language specific. It might be related to the fact that the antecedent acts as a discourse antecedent. Cf.:
(i) *Gestern ist ein Frosch erschienen. Die Prinzessin hat ??den/diesen Frosch nicht für einen Prinz gehalten.*
'Yesterday a frog appeared. The princess didn't consider ??the/this frog a prince.'

[12] The difference between French and German is the obligatoriness of topicalizing the coreferent element (be it the pronoun or the epithet) in the latter whereas it is left in baseposition in the former language.

general and not just on topicalization of the d-pronoun.

3.1.5.2. CP and NP

A similar conclusion can be drawn from the following paradigm, similar to the one above except for the fact that we are dealing with topicalization of an NP and Left Dislocation of a sentential element which is 'construed' with this NP and thus coindexed. Again we find that the possibility for Left Dislocation is crucially parasitic on topicalization, however not necessarily topicalization of a pronoun. First consider the following example:

(22) *Daß sie es ihm versprochen hätte, das/dieses/DAS Argument ließ die Prinzessin nicht gelten.*
that she it him promised had, the/this/THE argument let the princess not hold
'The princess didn't let the argument go through, that she had promised it to him.'

Notice that Left Dislocation is also the only environment where stress on the determiner is possible without explicit Identification. So we have further confirmation of the claim that Left Dislocation is somehow the opposite of Identification. The example in (22) is also interesting because of the following reason. Remember that I have argued that a CP can act as an identifier for NP. (I have rejected a complement relation and therefore an argument relation between the NP and the CP.) This assumption has the consequence that the CP is really adjoined to NP (rather than in complement position of N). The possibility to left dislocate the relevant CPs provides further empirical evidence for the adjunction-analysis. If the CP would be licensed as an argument of the NP then the Left Dislocation configuration would not be such that the topicalized and the left dislocated constituent are coindexed. The left dislocated constituent would rather be connected to a position embedded within the topicalized element (i.e. the complement position within NP). However, such a configuration is in general not possible for Left Dislocation as shown in the following example:

(23) **Den Mann$_i$, [den Bruder [dessen]$_i$] ich gesehen habe.*
The man, the brother d-pron$_{GEN}$ I seen have

In the sentence above the left dislocated element is coindexed with a complement of the topicalized element. There is no way the left dislocated element could be licensed in this configuration. We will see more of these examples in section VII.4. Moreover these sentences are further confirmation for the claim that there is no direct dependency

relation between the left dislocated element and the topicalized constituent. If there were a direct relation, then would not be obvious why the sentence in (23) should be excluded.

3.1.5.3. Relative clauses

A similar conclusion can be drawn upon the following example:

(24) (a) *Den die Prinzessin nicht küssen wollte, ??der/dieser/DER Frosch verwandelte sich in einen Prinzen.*
whom the princess not kiss wanted, the/this/THE frog turned himself into a prince
'It was the frog, whom the princess didn't want to kiss, who turned into a prince'

(b) *Den die Prinzessin nicht küssen wollte, DER verwandelte sich in einen Prinzen*
whom the princess not kiss wanted, d-pron. turned himself into a prince

Here we find a relative clause in left dislocated position. The head NP is in topicalized position and plays the same role as the d-words in other cases of Left Dislocation. In fact it is also possible to have a d-word instead of a full lexical NP (24b). Notice that Left Dislocation is the only instance of a relative clause preceding the head NP in German.

Similar arguments as the ones discussed in the previous section apply here. If the relative clause would be basegenerated (and thus inherently licensed) in the complement position of N then the grammaticality of the sentence in (24) is not expected. On the other hand the analysis proposed here, namely that a relative clause is licensed as an identifier (i.e. in a position adjoined to the identifyee) then the possibility to left dislocate a relative clause is expected. Moreover the analysis predicts that the relative clause in left dislocated position can just be interpreted as being construed with the DP rather than the NP. This is a correct prediction: the sentence in (24) has just one interpretation, it is not ambiguous in the way other relative clauses are. It only induces the presupposition that there are other frogs that do not have the property denoted in the relative clause. The presupposition that there are other individuals that also have the property denoted in the relative clause is not available.

One might object that the relative clause in (24) is a *free relative* and therefore

not really construed with the topicalized element. This proposal is easy to reject since headless relatives in German occur with a wh-pronoun instead of the d-pronoun:

(25) *Nimm, was/*das du willst!*
take what/d-pron. you want
'Take whatever you want!'

However the equivalent of (24) with a wh-pronoun in the relative clause is ungrammatical:

(26) **Wen die Prinzessin nicht küssen wollte, der/dieser/DER Frosch verwandelte sich in einen Prinzen.*
whom the princess not kiss wanted, the/this/THE frog turned himself into a prince

Again we are faced with the result that deleting the topicalized element would mean to delete a selected argument, violating the projection principle as well as the constraint on recoverability of deletion.

The conclusion of this section is that we cannot solve the 'intermediate binding problem' (i.e. that the trace would not be locally bound by an A'-Operator if there is an intermediate element which is coindexed but not an operator) like this problem is solved for intermediate traces. Under the assumption that left dislocated constituents are licensed as operators we are left with the solution presented here, namely that there is no chain between POp and TOp and no problem arises. The second conclusion is that - as our analysis predicts - Left Dislocation is parasitic on topicalization of any coreferent element (and not just on topicalization of a pronoun).

This analysis suggests that there is no real adjacency requirement between the left dislocated constituent and the coreferent element. The apparent adjacency requirement is simply a byproduct of the fact that the coreferent pronoun has to move to SpecCP, a position that is linearly adjacent to the position adjoined to CP. It is then expected that Left Dislocation can also be related to the variable in a long-distant way. The present analysis predicts that it is possible to find topicalization of the d-pronoun within an embedded clause and the left dislocated constituent in a position adjoined to the matrix clause. This is a correct prediction[13] as exemplified in the following example:

[13] Edwin Williams (p.c.) draw my attention to this prediction.

(27) (a) POp $[_{CP1}$.....$[_{CP2}$ TOp.....t]]
(b) *[Den Frosch]$_i$ [$_{CP}$ich glaube [$_{CP}$[den]$_i$ hat die Prinzessin t$_i$ geküßt.*
the frog [I believe [d-pron. has the princess kissed

POp is in a position to c-command the variable left behind by topicalization of the coreferent pronoun in the embedded clause (notice that it is crucial for the embedded clause to allow for V2). Notice that it is also possible to have the coreferent pronominal occurring in both clauses:

(28) *[Den Frosch]$_i$ [$_{CP}$[den]$_i$ glaube ich [$_{CP}$[den]$_i$ hat die Prinzessin t$_i$ geküßt*
the frog [d-pron believe I [d-pron has the princess kissed.

The possibility to have the coreferent d-pronoun occurring in both CPs is reminiscent of the possibility to have *partial wh-movement* as in the following sentence:

(29) *[Wen$_i$ glaubst du [wen$_i$ die Prinzessin t$_i$ geküßt hat?*
who believe you who the princess kissed has

Given the present analysis the existence of 'partial topicalization' w.r.t. Left Dislocation is not surprising. Finally notice that Left Dislocation is not really a root phenomenon. The present analysis predicts that it is only parasitic on topicalization. So whenever we find topicalization possible, then we expect Left Dislocation also to be possible. Again this is a correct prediction:

(30) (a) *%Den Frosch daß du geküßt hast, glaube ich dir nicht*
The frog that you kissed have, believe I you not
'As for the frog, I don't believe that you have kissed him'
(b) *%Den Frosch, den daß du geküßt hast, glaube ich dir nicht*
The frog d-pron. that you kissed have, believe I you not

My dialect does not show Doubly filled Comp filter - effects, i.e. it is possible to have a topicalized constituent preceding the complementizer.[14] In this case it is also possible

[14] There is one more thing to mention here. It is generally assumed that LD (as opposed to CLLD) is a root phenomenon. Now my analysis predicts that this is not really the case. It is just a phenomenon which is dependent on the possibility to topicalize. So in as much topicalization is a root phenomenon Left Dislocation is also. However, as (i) suggests this is not really true. Although I do not have anything to say on why the following sentence which is similar to (30) except for

to have a left dislocated constituent in this configuration. This piece of data is another piece of evidence that Left Dislocation is licensed by topicalization.

3.2. Empirical motivation

Besides the theoretical force for assuming that we are dealing with two independent chains there is also empirical evidence that suggests that this is really what we find.

The claim that there is no chain between POp and TOp forces us to say that the specific feature sharing between the left-dislocated constituent and the topicalized pronoun is not a matter of the relation between POp and TOp. The relation between these two elements is just one of discourse antecedence. However the two elements share more feature than other instances of discourse antecedence. Our analysis can explain this fact. In addition to being a discourse antecedent there is also a syntactic relation between the two elements, however this relation is a *mediated* one. Both, POp and TOp are linked to the trace in the matrix clause. Therefore we have to say that feature sharing is a byproduct of the fact that the two elements are indirectly related.

3.2.1. d-pronouns cannot have syntactic antecedents

A major empirical argument for the claim that the left dislocated element acts like a discourse antecedent for the topicalized pronoun is the fact that d-pronouns can never have syntactic antecedents. We have already seen that d-pronouns cannot have a bound variable interpretation in section VII.3.1.2. However it turns out that there is a more general restriction on d-pronouns. They cannot be syntactically bound by an antecedent within the clause at all. This is exemplified by the following sentences:

(31) (a) *Der Frosch hat gesagt, daß er/*der die Prinzessin küssen will*
The frog has said that he/*d-pron the princess kiss want
'The frog has said that he wants to kiss the princess.'
(b) *Die Prinzessin hat dem Frosch gesagt, daß sie ihn/*den nicht küssen will*

the fact that the complement CP occurs in extraposed position:
(i) **Ich glaube dir nicht, den Frosch (den) daß du geküßt hast*
I believe you not, the frog (d-pron) that you kissed have
So it seems that the possibility to have topicalization only in a complement clause that itself occurs in topicalized position is an instance of pied piping.

the princess has the frog said, that she him/d-pron. not kiss wants
'The princess said to the frog that she doesn't want to kiss him.'

Assuming that the d-pronoun in these cases is the same as the one used in case of Left Dislocation it is safe to conclude that the left dislocated constituent is really acting like a discourse antecedent rather than a syntactic antecedent for the d-pronoun.

3.2.2. Selectional restrictions are still available in Left Dislocation position

Given the assumption that the left dislocated constituent is analyzed as binding the variable in the clause we predict that certain selectional requirements have to be met in the position of the left dislocated constituent. Selectional requirements must be available in the head of a chain. This prediction is borne out as the following minimal pairs show:

(32) (a) *Eine Prinzessin küssen, das wollte der Frosch immer schon*
a princess kiss, d-pron wanted the frog always
'The frog always wanted to kiss a princess.'

(b) **Eine Prinzessin zu küssen, das wollte der Frosch immer schon*
a princess to kiss, d-pron wanted the frog always

(33) (a) **Eine Prinzessin küssen, das versprach der Frosch dem Heinrich.*
a princess kiss, d-pron promised the frog the H.
'The frog promised Heinrich, to kiss the princess.'

(b) *Eine Prinzessin zu küssen, das versprach der Frosch dem Heinrich.*
a princess to kiss, d-pron promised the frog the H.

In (32) the (modal) verb selects for a VP therefore the infinitival marker *zu* ('to') is not possible (since it can just occur within IPs)[15]. On the other hand the verb in (33) *versprechen* ('promise') selects for an IP (or CP) and therefore the infinitival marker *zu* is obligatory. If the only direct relation found in the representation above would be between the topicalized pronoun and the left dislocated constituent then the paradigm above would not follow straightforwardly. Notice that the topicalized pronoun is the same in both sentences, i.e. the neuter d-pronoun. Therefore it is not obvious how one could account for the fact that selectional restrictions are still available in left dislocated position.

[15] I will neither discuss the problem of the exact categorical nature of the constituent under consideration, nor the question as to where the infinitival marker *zu* is generated.

3.2.3. Matching and left dislocated constituents other than 'referential' NPs[16]

The second empirical support for the claim that there is a direct relation between the left dislocated constituent and the trace left behind via topicalization (rather than a relation between the left dislocated constituent and the topicalized element) involves Left Dislocation of some special NPs. It is usually assumed that (roughly speaking) specific reference is a prerequisite for Left Dislocation (cf. Altmann 1981):

(34) (a) **Irgendeine Prinzessin, die hat den Prinzen geküßt.*
Some princess or other, d-pron has the prince kissed

(b) **Ein Prinz, der ist normalerweise kein Frosch.*
a prince, d-pron is usually no frog

However, indefinite, non-referential NPs are possible candidates for Left Dislocation under certain circumstances. Consider the following paradigm:

(35) *Eine Prinzessin, das hat der Frosch gesucht*
a princess, d-pron$_{neuter}$ has the prince looked for

(36) (a) *Froschkönige, das haben wir gewählt*
frog-kings, d-pron$_{neuter,sg.}$ have we elected

(b) *Froschkönige, das sind Märchenfiguren*
frog-kings, d-pron$_{neuter,sg.}$ are fairy-tail-characters

(37) (a) *Zwei Männer, das sind einer zu viel*
two men, d-pron$_{neuter,sg}$ are one too much

(b) *Zwei Männer, das ist einer zu viel*
two men, d-pron$_{neuter,sg.}$ is one too much

In the cases above we find a left dislocated NP which is either feminine or plural. The pronominal element apparently does not agree in phi features. It invariantly occurs in the neuter singular form *'das'*, no matter whether the left dislocated NP is feminine (35) or plural (36). Moreover in (37) the verb can either agree with the left dislocated plural NP (37a) or with the singular d-pronoun (37b). Given my analysis this paradigm follows straightforwardly. The left dislocated element acts as an operator for the variable, i.e. it is in a chain relation with this variable. Rather than claiming that there is direct feature sharing between the two coreferent elements my analysis suggests that it is simply an effect of the relation between the trace and the operator. The paradigm

[16] I will not discuss what the actual generalization for the relevant NPs could be (i.e. I will not decide on whether we are dealing with simply non-referential NPs or non-extensional NPs or predicative NPs).

above shows that this is the right prediction. There is no matching in phi features between the coreferent element and the left dislocated constituent but there is a matching between the selectional requirements and the left dislocated constituent.[17]

We now have theoretical and empirical support for the claim that there is no direct (dependency) relation between the left dislocated constituent and the coreferent element. Moreover the empirical support for the present analysis is also important in the following respect. It shows that properties (i.e. VPs, IPs as well as predicative NPs) can be associated with discourse referents. Remember that we also have said that properties do not have any phi-features (cf. section III.5.), therefore the invariant occurrence of a neuter pronominal is expected. Finally the analysis presented here also supports the claim that discourse antecedence and syntactic antecedence are two independent phenomena, they can but need not go together. Left Dislocation is one instance where we have properties of both.

4. Left Dislocation vs. relative clauses

There is another (syntactic) issue to be addressed. It has to do with the difference between relative clauses and Left Dislocation. Let me start with a claim found in the literature on Left Dislocation: It has been argued that Left Dislocation in German triggers V2 (i.e. movement of the inflected verb to C°, cf. Zwart 1993). However, it seems to me that nothing in principle forces the verb to move to C° in case a constituent is left-adjoined to a CP.[18] Consider the output of a sentence construed with

[17] I think that the reason for the lack of phi-features (and neuter can be interpreted as a lack of phi features has to do with the fact that the discourse referent is really a property. Properties in general are referred to with neuter pronominals as can be seen in the following examples:

(i) *Grün, das war der Frosch hinter den Ohren.*
green, d-pron$_{neuter}$ was the frog behind the ears

(ii) *Der Frosch war [grün], hinter einem Ohr, aber der Prinz war [das], hinter beiden.*
the frog was green behind one ear, but the prince was that behind both.

In (i) we find a left dislocated adjective denoting a property. The coreferent pronoun appears in neuter. The same is true for intersentential anaphoricity: the adjective in the first sentence is coreferent with a neuter pronoun in the second sentence. If a (referential) definite DPs acts as the discourse referent a coreferent pronoun has to match gender

[18] Notice that in case of relative clauses it is not really true that a constituent is left-adjoined to a CP. Rather the CP is right-adjoined to the NP. However there is a striking similarity between the two configurations:

(i) NP (ii) CP
NP CP NP CP

In both cases (i)-(ii) we find two constituents that form a constituent. In both cases the NP

a preceding NP with and without movement of the verb:

(38) (a) *Den Frosch, den hat die Prinzessin nicht geküßt*
the frog, d-pron. has the princess not kissed
(b) *den Frosch, den die Prinzessin nicht geküßt hat*
the frog, whom the princess not kissed has

Both cases result in a wellformed output no matter whether the verb has moved (38a) or not (38b) the only difference being the interpretation. The sentence in (38a) is interpreted as an instance of Left Dislocation. In (38b) where the verb has not undergone movement we are dealing with an NP construed with a relative clause. The difference might be due to the fact that a V2 sentence is a fully saturated clause and is therefore interpreted as an assertion about the left-adjoined NP. On the other hand a relative clause (where the Verb is not moved to C°) is interpreted as a property that is predicated of the NP it is construed with. If one would stick to the claim that movement of a d-pronoun to SpecCP triggers verb movement one has to explain why this is not so in case of relative clauses. However, this would end up in a construction specific statement.

Our analysis makes an interesting prediction concerning another difference between relative clauses and Left Dislocation. This has to do with the licensing requirement of the adjoined NP. The 'head NP' of a relative clause is independently licensed within the clause it occurs in (as an argument) therefore it need not be licensed via the trace of the moved pronoun. This explains the following paradigms. A relative clause can be construed with an NP which is embedded in another constituent whereas Left Dislocation is just possible if the whole left-adjoined NP is coreferent with the pronoun in SpecCP (as we have already seen in example (23)):

(39) (a) *[Einer der Prinzen]$_i$, der$_i$ hat die Prinzessin geküßt*
one of the princes, d-pron.$_i$ has the princess kissed
(b) **[Einer [der Prinzen]$_j$], die$_j$ haben die Prinzessin geküßt*
one of the princes, d-pron.$_{pl}$ have the princess kissed

(40) (a) *[Einer der Prinzen]$_i$, der$_i$, die Prinzessin geküßt hat, ist gekommen.*
one of the princes, who.$_i$ the princess kissed has, has come

precedes the CP. The two configurations in (i) and (ii) can be interpreted as to saying that they differ as to which of the two constituents under considerations projects, i.e. which categorical feature percolates to the node created via adjunction. In (i) it is the adjoined phrase that projects further whereas in (ii) the target itself projects. If such an analysis is on the right track one still has to find a restrictive theory for determining which of the two nodes under consideration can project. I will leave this possibility as a matter of future research.

(b) *[Einer [der Prinzen]$_j$], die$_j$ die Prinzessin geküßt haben, ist gekommen.*
one of the princes, who$_{pl}$ the princess kissed have has come

In (39) the adjoined NP can just be licensed if it is coreferent with the topicalized d-pronoun, it is not enough if a subconstituent of the adjoined NP is licensed. On the other hand the NP in (40) is licensed independently (as subject of the matrix clause) and therefore it is possible to construe either a subconstituent of the NP or the whole NP with the relative clause. Remember that the relative clause is licensed as an identifier, i.e. it has to obey the LCI[19] (as well as PCI). Thus it is expected that the relative clause can be construed with a constituent more deeply embedded than the actual constituent the relative clause is adjoined to. The same is true in case the coreferent pronoun is more deeply embedded in another constituent:

(41) **[Die Prinzessin]$_i$, [[deren]$_i$ Vater] war sehr streng*[20]
the princess, whose father was very tough

(42) *[Die Prinzessin]$_i$, [[deren]$_i$ Vater] sehr streng war, hat sich durchgesetzt.*
the princess, whose father was very tough has won

Again the head of the relative clause is licensed independently as subject of the matrix clause whereas the NP in case of Left Dislocation must be licensed through the trace of the moved pronoun which is not the case if the coreferent pronoun is more deeply embedded in the moved constituent. Notice that here we are confronted with an apparent problem. This problem concerns coindexation of the relative clause as a whole and the head NP (which is crucial for the analysis of Identification, otherwise the PCI could not be derived). I have argued that the index of the whole clause is crucially dependent on the relative pronoun. The index of the relative pronoun percolates up to the dominating CP. In (42) the relative pronoun occupies the Specifier of the Specifier of the Relative Clause. However, I do not think that this is a particular problem for the analysis suggested here. Whatever has to be said about pied piping and how feature percolation within pied piped constituents takes place can apply to the analysis suggested here.

[19] Notice that according the LCI the relative clause in (35b) can either be adjoined to the embedded DP (*der Prinzen*) or alternatively to the entire DP (*einer der Prinzen*). 'DP-internal'-extraposition is not detectable on basis of this example. For the present argumentation the relevant representation does not really matter.

[20] It has to be noticed that this example is wellformed with a heavy intonation break, thus I assume that in this case it is interpreted as Hanging Topic. In fact it shows all other properties of Hanging Topic. A subject I will discuss in the next section.

5. Left Dislocation vs. Hanging Topic

Finally I want to briefly mention an implication the analysis of Left Dislocation suggested here has for Hanging Topics. I have argued that left dislocated constituents are only licensed if they can be mapped to a variable within the clause. Moreover we have seen independent evidence that the d-pronoun cannot act as a bound variable. However, it is sometimes claimed that topicalization of the coreferent pronoun is not really obligatory in German. A sentence like the following one is grammatical:

(43) *Der Frosch, die Prinzessin hat ihn/den nicht geküßt.*
the frog, the princess has him/d-pron. not kissed.

It can be shown that the sentence above has different properties than the instances of Left Dislocation we have been discussing so far. This construction is best analyzed as *Hanging Topic* rather than Left Dislocation. A crucial difference is for example the intonation pattern we find. Hanging Topic has comma intonation whereas Left Dislocation shows up with a single contour. Moreover, Hanging Topics do not have to match with the coreferent pronoun in all the features (i.e. there is no matching in case as exemplified in (43)). It has been noticed in van Riemsdijk & Zwarts (1994) that there are certain differences w.r.t. the context Left Dislocation and Hanging Topic can occur in respectively:

(44) Q: *Magst du den Peter?*
'Do you like Peter?'
A: *Den Peter, nee ich mag ihn nicht besonders.*
the Peter, no I like him not particularly

(45) Q: *Was ist eigentlich dein Verhältnis zu diesen Leuten?*
What is by the way your relation to these people
A: **Na ja, den Peter, ich mag ihn eigentlich nicht besonders.*
Well, the Peter, I like him actually not particularly
(van Riemsdijk & Zwart, 1994)

The different properties of Hanging Topic and Left Dislocation have to be explained somehow. I cannot really give a positive answer as to what licenses the Hanging Topic and what relation between the Hanging Topic and the coreferent pronoun in the matrix clause is. What seems to be at play here is that the Hanging Topic is a repetition of a previously mentioned NP. I do not think that we find an instance of binding the pronoun in the matrix (that could in principle act as a bound variable). There are several reasons that we do not want to assume such a relation. First the two elements only need to share their phi-features (like any other discourse antecedent). They need

not match in their case-feature. Secondly, a d-pronoun can also occur (as exemplified in (43)) and d-pronouns cannot act as bound variables. Thirdly it is also possible to have either the coreferent Hanging Topic (46a) or the pronominal element (46b) embedded in another constituent:

(46) (a) *...ach ja, [der schönste [der Prinzen]$_i$], die Prinzessin hat sie/die$_i$ übrigens gestern alle verschmäht.*
well, [the prettiest [the$_{GEN}$ princes]], the princess has them/d-pron yesterday all rejected

(b) *[Die Prinzessin]$_i$, der Frosch hat [ihren/deren$_i$ Vater] überzeugt*
[the princess]$_i$ the frog has [[her/d-pron]$_i$ father] convinced

If the Hanging Topic would be licensed by being directly mapped to the coreferent pronoun both sentences in (46) were expected to be ill-formed (as the parallel cases of Left Dislocation are). Finally, there is another piece of evidence that the Hanging Topic does not function as an Operator binding a variable. Consider the contrast below:

(47) (a) *Krieg und Frieden, das würde ich seinem Autor niemals zu lesen geben aber Dr. Schiwago, das schon*
War and Piece, d-pron. would I his author never to read give but Dr. Schiwago I pron. I would

(b) *Krieg und Frieden, ich würde das seinem Autor niemals zu lesen geben aber Dr. Schiwago schon.*
War and Piece, I would d-pron. his author never to read give but Dr. Schiwago I would

Sentence (47a) involving is ambiguous between the strict vs. the sloppy identity reading. This suggests that we are really dealing with an operator binding a variable. On the other hand the strict identity reading is the only one available in case of (47b) involving a Hanging Topic. This suggests that the non-topicalized pronoun as well as the possessive pronoun cannot be interpreted as bound variables. The exact nature of the licensing requirement for the Hanging Topic still needs an explanation.[21]

[21] However, it might be the case that English sentences referred to as Left Dislocation (as in *John, Mary really likes him* are also instances of Hanging Topics. The various properties still need to be tested.

6. The Presuppositions induced by Left Dislocation

Having established the syntax of Left Dislocation and thus having solved the apparent problem Left Dislocation seems to create for PCI I will now turn to the presuppositions induced by Left Dislocation. We will see that w.r.t. presuppositions Left Dislocation relates to topicalization in the same way as extraposed clauses with a correlate in the matrix relate to extraposed clauses without a correlate. The presence of the pronominal will be shown to induce a difference in meaning - confirming the Optionality Hypothesis. Moreover the following discussion will provide independent evidence for the adequacy of the representation of (pro)nominals suggested here.

Let us start with a comparison of Left Dislocation and topicalization[22]. Consider the following sentences:

(48) (a) *Der Hans, der hat es verstanden*
the John d-pron has it understood

(b) *Der Hans hat es verstanden*
the John has it understood

The two sentences differ minimally in that (48a) is an instance of Left Dislocation i.e. there is a DP adjoined to CP and a coreferent topicalized pronominal in the matrix. In (48b) we find an instance of topicalization, i.e. the lexical DP itself is topicalized (i.e. in Spec CP). Thus the presence of the pronominal seems somehow optional. According to what we have said about optionality in general we expect the following pattern. If an element is syntactically optional (like the pronominal above seems to be) then it must obligatorily be represented at some level of representation. We expect sentence (48a) to be obligatorily related to a certain interpretation. This interpretation should however be optional for the sentence in (48b). As a matter of fact it has been independently argued that there is indeed a special interpretation associated with Left Dislocation (of Conditionals, to which we will return in chapter VIII.) by Iatridou (1994). However, her analysis ends up with a problem, but one which can be solved by means of the analysis presented here. I will first briefly discuss her analysis and the problem she ends up with.

[22] The terms 'Left Dislocation' vs. 'topicalization' will be used to refer to sentences with or without a d-pronoun respectively. I will not discuss the adequate representation of 'topicalization', i.e. whether the relevant constituent occupies SpecCP or a position adjoined to CP cf. Koster (1978). For ease of exposition I will continue to assume that topicalization is always movement to SpecCP. For the present discussion however this is not crucial.

6.1. Iatridou (1994) on Conditionals

Iatridou (1994) argues that the presence of *then* in a conditional gives rise to a presupposition, in addition to the assertion of a statement, which conditionals without *then* are lacking. The presupposition Iatridou detects is as follows:

(49) Statement: if p, then q
Assertion: O[p] q
Presupposition: ¬O [¬p]q

Here *O* is the Operator induced by the conditional. In addition to the assertion that *if p then q* there is the following presupposition: it is not the case that if not p then (still) q. According to this the presupposition for a sentence like (50) can be paraphrased as (50b), where the presupposition is formulated positively, i.e. in terms of existence:

(50) (a) *If it's sunny, then Michael takes the dog to Pastorius Park*
(b) There are some cases in which it isn't sunny and in which Michael doesn't take the dog to Pastorius Park.

Iatridou concludes that conditionals with and without *then* are not to be equated, contrary to standard assumptions. This amounts to saying that the pronominal cannot be deleted. By means of this analysis Iatridou (1994) can account for several differences between conditionals with and without *then*. Yet, there remains one problem: why is it that *only if*- and *even if*-conditionals block the possibility to have *then* although the presupposition in (49) would be compatible with the interpretation of these focus particles? To account for this fact Iatridou proposes the possibility to assimilate the relationship between an *if*-clause and *then* to (German and Dutch) Left Dislocation[23], which shows the same pattern: *only NP* and *even NP* are not possible in left dislocated position. Whatever is responsible for this phenomenon can also be at play in Conditionals that are instances of Left Dislocation. Notice that Iatridou still does not give an account for this fact. We will see later that matters are not as simple as they are stated here:

(51) (a) **Only if it is sunny, then Michael takes the dog to Pastorious Park*
(b) **Even if it is sunny, then Michael takes the dog to Pastorious Park*

[23] Notice that this analysis, i.e. to treat these instances of conditionals on par with Left Dislocation is further confirmed in chapter VIII., where I will look at conditionals under the light of the analysis for Identification presented here.

(52) (a) **Nur der Hans der ist gekommen.*
Only the J. d-pron is come
(b) **Sogar der Hans der ist gekommen.*
even the J. d-pron is come

Assuming that Conditionals with *then* are instances of Left Dislocation to account for the *only/even* problem[24] it is of course expected that Left Dislocation of NPs carries a similar presupposition. Given the logic of her argumentation the presupposition of the sentence in (53a) should be the one given in (53b):

(53) (a) *Der Hans, der hat es verstanden.*
the John, d-pron has it understood.

(b)

Statement:	NP_i der, P
Assertion:	NP_i P
Presupposition:	There is some NP_j (NP_j is not NP_i) such that ¬P is true of NP_j, i.e. [NP_j ¬P].

In (53b) NP_i is the left dislocated NP. If the property P is asserted about NP then the presence of the pronoun gives rise to the presupposition that there is some other discourse referent NP_j who does not have this property P. So the sentence in (53a) is necessarily associated with the following interpretation (as opposed to instances of topicalization where this interpretation is just optional):

(54) *There is somebody other than Hans who did not understand it.*

There is one apparent problem for the assumption that Left Dislocation is necessarily associated with such a presupposition, which I will discuss later. For the moment let us assume that the presupposition above is indeed what we find. Iatridou does not address the question why the presence of the pronominal triggers the presupposition. I think that the analysis suggested here makes it possible to derive the presupposition. I will argue that it follows from the representation of (pro)nominals I have introduced in chapter III.

[24] Notice that in her article Iatridou discusses several possibilities for solving this problem, the Left Dislocation analysis being just one of them. I think that the present discussion crucially supports this analysis.

6.2. Deriving the Presupposition

Recall that I assume that every nominal is associated with two indices. So far I have only sketched the representation of names. Without really going into detail I will assume that their representation is similar to the ones of pronominals. In case of names X denotes the set of possible individuals. An indirect piece of evidence stems from the fact that contrastive stress on proper names induces a set of alternatives (in the sense of Rooth (1985)). We can assume that the set of alternatives is provided by X. Consider a sentence like (55) and the representation associated with the stressed name:

(55) (a) *I met JOHN (and not Mary).*
(b) X = set of possible individuals
x = JOHN
$x \subseteq X$[25]

By means of the subset relation the set of alternatives is established. Under normal circumstances it is always the case the discourse referent associated with X is not accessible. I have argued that the presence of x (in case of common nouns provided by the determiner) prohibits access to X. As the sentence in (55) above indicates (and as we have seen in section IV.7.) contrastive stress on N makes X available. This availability corresponds to the presupposition we arrive at, i.e. the set of possible antecedents (which is the set of possible individuals) is 're-introduced' inducing the presupposition *that there are other individuals I did not see*. Notice that this assumption is reminiscent of Ward's (1985) analysis of focus. He assumes that a proper name (in many cases) comes with two denotations. One corresponds to the actual discourse referent the other corresponds to some sort of *scale*'[26]. If a name is focussed this scale corresponds to the set of alternatives which is at the same time the focus value of the nominal. In our terms the second denotation corresponds to X that denotes the set of possible antecedents. We can assume that this set of possible antecedents is ordered w.r.t. the actual likelihood (i.e. possibility) to occur as the intended antecedent resulting in a scale-interpretation. I will come back to the relevance of a scale being

[25] Notice that I am still assuming a subset relation $x \subseteq X$. It might seem more appropriate to have $x \in X$ instead. In case of names the denotation of x will mostly correspond to a singleton set. For proper nouns and pronominals I have assumed $x \subseteq X$ in order to account for plural NPs. Since it is also possible to have plural names as in *Maiers sind gekommen* (Maiers have come) I will continue to use the subset-of relation rather than the member-of relation.

[26] Thus a sentence like
(i) *Fido$_i$ they named it t$_i$*
is associated with an interpretation that *Fido* is a very unlikely name for the dog to be named.

associated with X in the course of this section and more explicitly in chapter VIII.

I would now like to argue that besides stress there is another possibility to make the set of possible antecedents available, namely by means of the pronoun. Remember that I have argued that in case of extraposition the pronoun is associated with an index X (associated with the set of possible antecedents) and an index x which still needs Identification. Also, remember the presupposition associated with any restrictive relative clause, namely that there is at least one individual having the property denoted by N which does not have the property denoted by the relative clause. This is however rather similar to the presupposition Iatridou (1994) argues to be associated with Left Dislocation. Consider in this light the following representation for the nominal elements involved in Left Dislocation:[27]

(56) (a) *Der Peter, der ist gekommen.*
the Peter, d-pron. is come

(b) DP: x = Peter
X = set of possible individuals[28]
Pron: X = set of possible (male) antecedents
x = Peter
$x \subseteq X$

Since the presence of the pronominal is syntactically optional we expect that in order to be licensed it needs to be present at some level of representation. However x itself cannot be the crucial factor since it is already present at domain D (by means of the left dislocated DP). According to this line of reasoning it must be X which is crucial for licensing the presence of the pronominal. This means that X has to be represented at domain D. So I would like to conclude that the pronominal reintroduces the set of possible antecedents. Its presence makes this set available again. In a way this is a

[27] I have abbreviated the representation in the already familiar way: the pronominal would have to be associated with an index y, which is later equated with the index of the DP.

[28] Notice that there is a contrast between Left Dislocation of a 'bare-name' and Left Dislocation of a name that comes with a determiner, the former being degraded.
(i) *?Peter, der ist gekommen*
Peter d-pron is come
(ii) *Der Peter, der ist gekommen.*
the Peter, d-pron is come
This contrast might suggest that a name in German is associated with X only in case the determiner is present. Or alternatively that this X is more easily available if the name still occurs within N rather than D.

similar phenomenon as contrastive stress (in which case intonation is the relevant factor) to make X available at domain D. The individual denoted by DP is contrasted to a set of alternatives (provided by X). Notice that in the literature on Left Dislocation in German (and Dutch) this construction is often referred to as *Contrastive* Left Dislocation. Under the present analysis the contrastiveness reduces to the same phenomenon in both cases. The denotation of NP (X) is rendered available either by means of stress or by means of an additional pronominal. The denotation of X together with the fact that x has to be a subset of X derives the presupposition detected in Iatridou (1994).

Now it should also be clear how Left Dislocation is really the opposite of Identification. In case of Identification we first establish a set of possible antecedents, the individual to be picked out still has to be identified. In case of Left Dislocation the discourse referent is already established and also available (in fact it has to be the topic, which is somehow the most available discourse referent). However, here the presence of the pronominal (re-)introduces a set of individuals (of which the discourse referent of DP is a member). It seems to me that the presupposition of Left Dislocation provides indirect support for the representation of nominals I am assuming. Furthermore it also provides independent evidence of the analysis of apparently optional insertion, i.e. for the Optionality Hypothesis.

There are two more effects we can derive by this analysis. First the analysis provides an explanation for the fact that Wasow's (1972) Novelty Condition is obeyed in case of Left Dislocation. The coreferent topicalized nominal element is only licensed if it provides a superset for the left dislocated nominal. The difference between having a pronominal or a full lexical NP in SpecCP is simply the nature of the superset which is re-introduced. In case of a pronominal it is just the set of possible antecedents (with no further inherent restrictions). In case of a full lexical NP a further restriction is added, namely that it has to be a member of the set of individuals satisfying the property denoted by N. In that case the superset itself is restricted - it cannot be any possible antecedent. A question that might arise is why (extraposed) identifiers do not violate the Novelty Condition. Notice that I explicitly argued that Identification is not an instance of an anaphoric relation. It is really the opposite. One of the crucial properties of the identifyee is that it does not have enough descriptive content to pick out the intended referent. Therefore it is an inherent property of Identification that it adds further restriction, it must contain more descriptive content than the identifyee itself. Again this behavior provides further confirmation for the claim that a cataphoric relation is the opposite of an anaphoric relation.

The second prediction has to do with the problem of optionality. Remember that we analyzed optionality by means of a special notion of licensing. We have seen in section III.3. that syntactic optional phenomena have to be licensed at domain D. On the other hand in case a given phenomenon is not syntactically optional then an optional interpretation arises: i.e. the interpretation associated with a given representation can but need not arise. For the present cases we predict that if the coreferent element is syntactically obligatory (i.e. syntactically licensed) then it need not have a representation at domain D. Therefore we expect that the contrastive interpretation can but need not arise. We have already seen instances where the presence of the coreferent element is necessary due to syntactic reasons. One relevant example is repeated here for convenience and the prediction is indeed borne out. The sentence above is felicitous even if there is no superset (of possible antecedents) available - thus the presupposition does not necessarily arise.

(57) *Daß der Frosch ein Prinz ist *(dessen) war sich die Prinzessin nicht bewußt.*
that the frog is a prince *(d-Pron) was herself the princess not aware
'The princess was not aware of the fact that the frog was a prince.'

Now I will provide some empirical evidence for the present claims. I will discuss some contexts that can be taken as an empirical test for the presupposition. In doing so we will discuss the problem Iatridou's analysis seems to be faced with - as mentioned at the beginning of this section.

We have seen that (contrastive) Left Dislocation presupposes that there is at least some other individual for whom the assertion in the CP is not true. This means that the discourse referent of the left dislocated DP may not exhaust the particular that is asserted to be true for this denoted referent. In other words, Left Dislocation implies non-exhaustiveness, the superset may not be equated with the subset. Iatridou (1994) gives the following example (due to Irene Heim) that seems to falsify this prediction.

(58) *Alle haben die Vorlesung verstanden. Hans hat sie verstanden. Maria hat sie verstanden. Und unser Freund Peter, der hat sie auch verstanden.*
Everybody has understood the lecture. J. has understood it. Mary has understood it. And our friend Peter, der has understood it as well.

(58) seems to create a problem. The given discourse makes it explicit that the intended superset contains only individuals for whom the assertion is true. We would expect that Left Dislocation is not possible, since the discourse seems to imply that there is no individual for whom the asserted property is not true. However, as we will see, in this

case we find is another presupposition arising. For the moment let us abstract away from this second presupposition that is responsible for the well-formedness of (58) and let us assume that Left Dislocation here is indeed not well-formed. This would then confirm the prediction that exhausting the set of entities for which a certain property holds results in ill-formedness of Left Dislocation. The following discourse, which slightly differs from the one given in Iatridou further confirms this prediction. (Notice that the difference between (58) and (59) is crucial as I will discuss below).

(59) *Alle haben die Vorlesung verstanden. Der Peter, (#der) hat sie zum Beispiel verstanden.*
Everyone has the lecture understood. For instance the Peter, d-pron has it understood

With normal intonation the discourse above is not wellformed if the d-pronoun is present. In case the d-pronoun is syntactically forced we expect that no such ill-formedness arises because the pronominal can be expletive w.r.t. domain D - thus exhaustiveness is not implied:[29]

(60) *Die Prinzessin war sich überhaupt nichts bewußt. Daß der Frosch ein Prinz ist dessen war sich die Prinzessin zum Beispiel nicht bewußt.*
The princess was not aware of anything. That the frog is a prince (ProPP) was herself the princess not aware
'The princess was not aware of the fact that the frog was a prince.'

Let me now discuss Iatridou's (1994) apparent problem induced by the well-formedness of the example in (58) more carefully since it is indeed true that there is a well-formed interpretation available.

[29] The presence of the head of the relative clause behaves differently in the context of Left Dislocation. Although it is not syntactically optional it still forces the presupposition. This follows since full lexical NPs are necessarily associated with a representation at domain D. Therefore the following discourse is not well-formed:
(i) *#Alle Frösche haben sich ein einen Prinzen verwandelt. Den die Prinzessin nicht küssen wollte, DER Frosch hat sich zum Beispiel in einen solchen verwandelt.*
All frogs have turned into a prince. Whom the princess didn't want to kiss. THE frog has for example turned into one.

6.3. Another Presupposition

Let us look at (58) in more detail. The line of argumentation is as follows. The first sentence, which includes universal quantification over individuals who understood the lecture, is intended to exclude the possibility that there is at least one individual who didn't understand it. This excluded possibility however corresponds to the presupposition Left Dislocation should carry. Therefore the sentence involving Left Dislocation is not expected to be good in this context. I will now argue that the sentence under consideration carries a different presupposition, which is in general available for Left Dislocation. This presupposition will however turn out to be compatible with the context in (58) and therefore this example does not cause a problem anymore. Moreover it will turn out that the two presuppositions correspond to two different intonation patterns. Consider the relevant sentence (as it occurs in (58)) repeated here in isolation and the presupposition it is associated with:

(61) (a) *'..und unser Freund der Peter, der hat es auch verstanden'*
and our friend the Peter d-pron has it as well understood

(b)

Statement:	NP_i der_i P
Assertion:	NP_i P
Presupposition:	(usually) NP_i such that ¬P is true of NP_i), i.e. [NP_i ¬P].

As indicated in (61b) the presupposition does not involve any other discourse referent. It simply says that usually the individual under consideration does not have the property that is asserted about him in this clause. We can paraphrase the presupposition as follows:

(62) It was highly unexpected for Peter to understand the problem. Usually he does not understand anything at all.

Notice that this interpretation is forced in (62) due to the presence of *auch* ('as well'). If this presupposition is chosen then the presupposition Iatridou (1994) predicts for Left Dislocation is not present anymore. Notice also that the two different presuppositions available for Left Dislocation correspond to two different intonation patterns[30]:

[30] The two different intonation patterns have been discussed by Hanneke van Hoof at the Tilburg Workshop on Left Dislocation 1994.

(63) (a) *und der HANS/ der hat es diesmal auch verstanden.*
and the J. d-pron has it this time as well understood
(b) *Der Hans\ DER hat es verstanden.*
the J. d-pron has it understood

In (63a) the left dislocated constituent has a rising, non-nuclear accent and the d-pronoun is deaccented whereas in (63b) the leftdislocated constituent has nuclear accent and the d-pronoun has non-nuclear rising accent. Notice that according to what we have said so far this intonation pattern is not unexpected. In case the left dislocated noun is stressed, X is already available, since stress on N is a device to make X available. Therefore it is expected that the d-pronoun does not make X (i.e. the set of alternatives) available[31], since it is not crucial anymore for licensing the presence of the pronoun.

We can now return to the problem of *only* and *even* as discussed at the beginning of this section. It will turn out that both of these focus particles are possible in Left Dislocation. However, *only* forces the first presupposition, whereas *even* forces the second one. This means that either of the two elements are restricted to one of the two possible environments. Consider again the relevant sentences that were taken to be ill-formed (following the judgements given in Iatridou 1994):

(64) (a) *(*)Nur der Hans, der hat es verstanden.*
Only the John d-pron has it understood
(b) *(*)Sogar der Hans, der hat es verstanden.*
Even the John, d-pron has it understood

As Iatridou mentions (citing Irene Heim p.c. for this observation) (64a) is a possible sentence, but only in an environment that can be glossed as (65):

(65) The material was so difficult that hardly anybody understood it. Only Hans der understood it.

'nur' ('only') forces the interpretation that it is highly implausible that anybody can understand the lecture. Therefore it is compatible with the presupposition that there are other people who do not understand the lecture. As expected we can just have the

[31] As a matter of fact in this case the presupposition does not seem to be dependent on the pronominal since the same sentence without the pronominal does have the same presupposition:
(i) *und der HANS hat es diesmal auch verstanden.*
and the JOHN has it this time as well understood.

Contrastive Intonation pattern, with stress on the d-pronoun:

(66) (a) *Nur der Hans, DER hat es verstanden.*
only the John d-pron has it understood
(b) **Nur der HANS, der hat es verstanden.*

Furthermore Iatridou mentions that there is no such effect for (64b). The sentence is still ill-formed in the following environment:

(67) The material was so easy that every student understood it. ?*Even Hans der understood it.

However, under the following interpretation Left Dislocation of a constituent containing *even* is rendered well-formed:

(68) Hans is the very least person to understand anything about philosophy but the lecture yesterday was so easy. Even Hans, der understood it (this time).

As we have seen above there is another presupposition available for Left Dislocation. It concerns the likelihood of the individual under consideration to have the property asserted about it. As we have seen it results in an interpretation where the discourse referent is highly implausible, to have this property. This is exactly what we find to be the environment given in (68). *Sogar* ('even') forces this interpretation so we expect it to be well-formed in this case of Left Dislocation. We have seen that the two presuppositions are associated with two different intonation patterns. Since (68) indicates that *sogar* is compatible with the second presupposition it is expected that it is only compatible with the second intonation pattern.

(69) (a) *Sogar der HANS, der hat es verstanden.*
even the John d-pron has it understood
(b) **Sogar der Hans, DER hat es verstanden.*

We now make the prediction that the same kind of presupposition is also available for *if p then q*-conditionals. Consider the following example discussed in Iatridou:

(70) *#If I were the richest linguist on earth, then I wouldn't be able to afford this house.*

According to her analysis this sentence is expected to be odd since the superlative induces an exhaustive antecedent and therefore it makes the presupposition of *then*

false. Given what the possibility for another presupposition (that is compatible with an exhaustive antecedent) we predict the sentence to be good in that case. Consider under this light the following example:

(71) *und wenn ich der REICHSTE Linguist auf Erden wäre, dann würde ich mir das Haus auch nicht leisten können.*
and if I would be the richest linguist on earth, then I couldn't afford it either

Like in the example involving Left Dislocation of a DP we find *und* ('and') introducing this statement and *auch* ('as well') in the matrix clause. Moreover we find the same intonation pattern: rising intonation on the antecedent and a deaccented d-pronoun. The presupposition we find here is the same as the one described above: it is highly unexpected that I would be the richest linguist on earth. It seems to be the case that the unexpected part of the conditional is the one which gets stressed.

Moreover we expect the same facts as in case of Left Dislocation concerning the possibility to have *only* or *even* respectively within conditionals. This prediction is again borne out (for *only* this has already been observed in Iatridou 1994):

(72) *John is such a homebody that he never leaves the house. Only if there is a heatwave, then he'll go out.*

For *even* we expect that the presupposition concerning the *implausibility* has to do with the antecedent. Remember that in case of Left Dislocation of a DP *even* was good in case the person denoted by the left dislocated DP is highly implausible to be able to do whatever is expressed in the CP:

(73) *Sogar wenn ich die REICHSTE Linguistin wäre, dann würde ich mir das Haus immer noch nicht leisten können.*
Even if I would be the richest linguist, then I could still not afford the house

Again we find the same intonation pattern. *nur* is compatible with stress on the d-pronoun, whereas *sogar* is only compatible with a deaccented d-pronoun:

(74) (a) *Nur wenn es eine Hitzewelle gibt, DANN geht Peter aus.*
only if there is a heatwave, THEN Peter goes out
(b) **Nur wenn es eine Hitzewelle gibt, dann geht Peter aus.*
only if there is a heatwave, then Peter goes out

(75) (a) *Sogar wenn ich die REICHSTE Linguistin wäre, dann würde ich mir das Haus auch nicht leisten können.*
Even if I would be the richest linguist, then I couldn't afford the house
(b) **Sogar wenn ich die reichste Linguistin wäre, DANN würde ich mir das Haus auch nicht leisten können.*
Even if I would be the richest linguist, THEN I couldn't afford the house

It is clear that in case of conditionals talking about an implausibility for something to happen is related to conjunctive (subjunctive) morphology. So the second kind of presupposition detected here is reminiscent of counterfactual conditionals.

It seems to me that the two different interpretations have something in common, namely the fact that there is a presupposition concerning the implausibility for a certain proposition to hold. This implausibility is again reminiscent of a scalar interpretation (as mentioned above). The difference between the two kinds of presuppositions discussed here is simply a matter of what constitutes the proposition which is presupposed to be unlikely to hold. Compare the following:

(76) (a) *und der Peter, der hat es auch verstanden*
and the Peter, d-pron has it as well understood
XP $|_{VP}$d-pron ...]
(b) *und wenn ich auch die reichste Linguistin wäre, dann würde ich mir das Haus auch nicht leisten können.*
and if I as well the richest linguist were, then would I myself the house as well not afford could
if XP $|_{VP}$ d-pron ...|

In both cases there has to be common knowledge about XP. So in (76a) in order for the sentence to be felicitous it has to be common knowledge that Peter is very unlikely in general to understand anything. Therefore the following two responses to an utterance of (76a) are possible, depending on the background of the hearer:

(77) *Why, do you say that?*
(a) *Is Peter such an idiot?*
(b) *Was the lecture so easy?*

In (77a) the hearer knws that the lecture was very easy. He can infer that Peter is not very smart. Whereas in (77b) the hearer knows that Peter is not very smart and so from the utterance in (76a) he can infer that the lecture was very easy. The same is true for

the (counterfactual) conditional in (76b): there has to be a common knowledge about the unlikelihood of XP (the antecedent) to be true. Again the following two responses are possible, depending on the background of the hearer:

(78) *Why do you say that?*
  (a) *Is it so unlikely that you (as a linguist) once become rich?*
  (b) *Is the house so expensive?*

If the hearer knows that the house is very expensive he can infer the unlikelihood of the antecedent to be true (78a). On the other hand if the hearer knows about the unlikelihood of the antecedent in the *if*-clause he can infer that the house must be very expensive. Compare these facts to the other presupposition:

(79) (a) *Der Hans, DER hat die Vorlesung verstanden*
  the J. d-pron has the lecture understood
  XP [$_{YP}$d-pron...]
  (b) *Wenn die Sonne scheint, DANN gehe ich gern spazieren.*
  If the sun is shining, then I like to go for a walk
  XP [$_{YP}$d-pron...]

Here in both cases there has to be a common knowledge about something which is related to the proposition expressed in YP. For (79a) the common knowledge corresponds to the fact that the lecture was very hard to follow. This utterance tells nothing about the abilities of Hans. But since it is so hard to understand the presupposition arises that there must be some other people who did not understand the lecture. Thus the following responses to (79a) are possible:

(80) *Why do you say that?*
  (a) *Was the lecture so difficult?*
  (b) *Is John so smart?*

If the hearer knows that John is very smart, he can infer from (79a) that the lecture must have been very difficult. On the other hand if the hearer knows that the lecture was very difficult he can infer that John is very smart. The same can be shown for (79b): there is a common knowledge about the content of YP, namely that in general it is the case that the speaker doesn't like to go for a walk although it is not excluded that he would ever go for a walk. Again there is no common knowledge necessary for the antecedent. Depending on the hearers background different responses are possible:

(81) *Why do you say that?*
  (a) *Don't you like to go for a walk in general?*
  (b) *Is the sun not shining very often here?*

Let me briefly summarize what we found in this section. There are two different presuppositions Left Dislocation can induce. Both have to do with the likelihood for a certain proposition to be true, i.e. both are scalar in nature. The first one (originally detected in Iatridou (1994) for Conditionals) presupposes that there are other individuals, circumstances, for which the proposition expressed in the 'matrix-clause' is not true. I have argued that this presupposition arises because the pronominal makes the discourse referent X (the set of possible antecedents) available. The second presupposition concerns the unlikelihood for the left-dislocated constituent to have a certain property (or in case of conditionals to be true). This suggests that there is something like a counterpart of counterfactuals in case of Left Dislocation of a nominal element. In that case we could argue that X of the pronominal is associated with the 'intension' of a proper name, i.e. the properties of a given individual.

Let me just briefly mention that Karttunen & Peters (1979) come to a similar conclusion w.r.t. the interpretation of *even* (which I already sketched in section VI.2.3.). They claim that *even* gives rise to two different implicatures (presuppositions in our terms): one *existential* and one *scalar* in nature. Thus consider their example with the associated interpretation:

(82) *Bill likes even MARY*
  Focus of *even*: *Mary*
  Scope of *even*: *Bill likes x*
  Existential implicature: *There are other x under consideration besides Mary such that Bill likes x.*
  Scalar implicature: *For all x under consideration besides Mary, the likelihood that Bill likes x is greater than the likelihood that Bill likes Mary.*
  (Karttunen & Peters, 1979: 26)

There would be many more interesting issues to be addressed. But this would go beyond the scope of the present discussion. What I have said in this chapter should suffice to capture the facts concerning Left Dislocation in the light of the analysis of Identification.

VIII. EXTENDING THE ANALYSIS: Other Cases of Identification

In this chapter we will extend the analysis discussed so far to some other related phenomena. I will argue that all instances of sentential elements that are construed with a correlate in the matrix clause are licensed as identifiers. The discussion in this chapter will mainly concern result clauses, comparatives and conditionals. In addition I will also mention adverbial clauses of cause and place. For reasons of space I will only discuss properties that are relevant in the light of my analysis. It is not the purpose of this chapter to give a fullfledged analysis of conditionals or result clauses. I only want to show how (some of) their properties can be derived by means of the analysis of Identification.

1. Result Clauses

Let us start with result clauses. Consider the following examples:

(1) (a) *Peter hat so viele Krügerl getrunken, daß er nicht mehr nach Hause kam.*
Peter has so many pints drunk, that he not anymore to home came.
'Peter has drunk so many pints that he couldn't go home anymore.'

(b) *Peter hat zu viele Krügerl getrunken, um noch nach Hause zu gehen.*
Peter has too many pints drunk, for still to home go
'Peter has drunk too many pints for him to be able to still go home.'

I will assume that the entire DP (containing *so*) is the identifyee. However, the presence of *so* has a crucial influence on the interpretation of this DP. Since this analysis differs from other proposals found in the literature I will briefly discuss an analysis which makes use of the *complement principle*[1] (rather than Identification), namely the one advocated in Gueron & May (1984) (henceforth G&M).

[1] I will discuss the 'complement'-principle in more detail in section Appendix I.

1.1. Result Clauses and the Complement Principle

It is in fact not a new proposal to treat result clauses and other instances of extraposed clauses (i.e. relative clauses) in a similar fashion. The syntactic properties of result clauses have been compared to relative clauses in various approaches. Here I will discuss the analysis given in G&M, since the main point I am going to argue against is advocated there in a very elaborate way. They observe that relative clause extraposition and result clause extraposition share the property that they involve extraposition from "headed" constructions.[2] In our terms this simply means that those are extraposed elements which have a correlate in the matrix, i.e. a clausal element identifying a (nominal) element in the matrix clause. G&M's main point w.r.t. this parallelism is as follows:

> "Result Clauses may be found as complements to *so* or *too* occurring as specifiers either of AP or of NP, co-occurrence in the latter case being required with *many* or *few* (for count nouns) or with *much* or *little* (for mass nouns); as complements to certain adverbs, like *sufficiently*, that can modify *many* or *few*; and as complements to certain quantity phrases, such es *enough*, *such a large*, and *just the right*, the latter two in cooccurrence with the noun *number*" (G&M, 1984: 1)

A crucial difference between their analysis and the one I am proposing concerns the nature of relation assumed to hold between the two elements under consideration: G&M claim that the relation between the two elements can be characterized as a complement relation whereas I argue that it is an instance of Identification. Furthermore, following other approaches to result clauses (cf. Liberman 1974, Rouveret 1978) they assume that whereas in relative clauses the head is the NP, in case of result clauses the head of the construction is the QP specifier of the whole phrase. Here we find another difference between standard treatments and the analysis I am advocating

[2] For convenience let me just summarize the difference between relative clause and result clause extraposition as they analyze it. I will discuss G&M's approach in more detail in section Appendix I.
Extraposition from NP and extraposition from QP (i.e. result clause extraposition) differ in
a) the landing site of the extraposed constituent:
extraposition from NP attaches to S, extraposition from QP attaches to S'
b) the landing site of QR of the head NP/QP in the matrix clause:
NP undergoes QR adjoining to S
QP undergoes QR adjoining to S'
It is assumed that the relevant landing site of extraposition derives from the complement principle (holding at LF) since the relevant head has to undergo QR and again the relevant heads differ in their landing site - the output however has to result in a configuration where the head governs the complement.

here: I will argue that the correlated element is not only *so* but the entire DP containing this (pronominal) element. Evidence for the former (i.e. G&M's) view is provided by the following empirical facts.
1) Dropping the QPs *so* and *too* in the relevant examples is not possible (3).
2) there seems to be a restriction on several features of the 'complement clause' depending on the occurrence of *so* or *too*, respectively (3):

(2) (a) *So many books have been published recently, that I haven't been able to read them all*
(b) *Too many books have been published recently, for me to be able to read them all*

(3) (a) **Many books have been published recently, that I haven't been able to read them all*
(b) **Many books have been published recently for me to be able to read them all*

(4) (a) **So many books have been published recently, for me to be able to read them all.*
(b) **Too many books have been published recently, that I haven't been able to read them all.*

What G&M intend to show with these examples is first that extraposed clauses are indeed complements to QP and secondly that there are dependency relations requiring *so* to be paired with a finite clause and *too* with an infinitive. Given these examples one could indeed conclude that the QP on its own (rather than the whole DP) is the head of the construction. The feature of the result clause that the quantifying element seems to determine is the (non-)finiteness of the extraposed clause as well as the nature of the complementizer. Notice however that G&M are not really explicit as to how this feature is determined. They only assume some sort of 'dependency'. Of course it cannot be denied that there is some sort of dependency, however I think that something more has to be said about the nature of this dependency relation. A relation that seems to show that a particular element determines some categorical properties of another element is generally viewed as involving c-selection. In this section I would first like to argue that we need not necessarily conclude that the restrictions on the nature of the result clause are determined by means of c-selection. Secondly I will show that assuming only the QP Specifier as the head of the relation rather than the entire DP induces some severe problems.

I think that the restrictions on various features of the result clause are dependent on some more general (semantic) properties. They need not necessarily be analyzed as selectional restrictions of the quantifying expression. This is not an unmotivated

assumption given the following facts:

(5) (a) *Ich weiß, daß/*?ob Peter gerne Bier trinkt*
I know, that/*?if Peter likes beer drinking
(b) *Ich weiß nicht, ob/*?daß Peter gerne Bier trinkt*
I know not, if/*?that Peter likes beer drinking
(c) *Ich wußte nicht, ob/daß Peter gerne Bier trinkt.*
I knew not, if/that Peter likes beer drinking.

These examples show that the choice of the complementizer cannot only be due to properties of the selecting verb. Negation as well as tense seem to have an influence on the choice of the complementizer. The following facts about German result clauses support this claim. It is true that the result clause without a correlated *so* (6a) sounds quite odd (similar to what is found in English). However, German has a complementizer that allows for a result reading without a QP correlate in the matrix (6b):

(6) (a) *#Ich habe viele Krügerl getrunken, daß mir schlecht wurde*
'I have drunk many pints drunk, that I got sick.'
(b) *Ich habe viele Krügerl getrunken, sodaß mir schlecht wurde.*
I have drunk many pints, so that I got sick'

In (6b) we find a complex complementizer consisting of *'so'* and *'daß'*. Here the result of getting sick is not directly related to a certain amount of beer but to the act of drinking (lots of) beer in general. Again according to my analysis this is not a surprising effect since I will show that any kind of property (and a VP is analysed as a property, i.e. a predicate) can induce a scalar interpretation, i.e. a certain amount of a given property can suffice to have the effect of a certain result. In the case above it is not even the quantifying determiner that is responsible for the amount reading. The following sentence containing a mass noun is also wellformed with a result clause:

(7) *Ich habe Bier getrunken, sodaß mir schlecht wurde.*
I have drunk beer, so that I got sick

This means that result clauses are not only possible with *'so'* but with any property. We will see some more examples supporting this claim in section VIII.1.2. I would like to conclude that the restriction on features of the result clause do not provide enough evidence to postulate a direct (selectional) relation between the QP and the result clause.

There is another problem with the view that *so* is the (c-selecting) head of the result clause. Notice that c-selection is normally instantiated in a very local way (i.e. under sisterhood). Thus it is a surprising fact that the two elements cannot occur adjacent to each other (at least not at the level of S-structure[3]).

(8) (a) **So that I haven't been able to read them all many books have been published recently*
(b) **Too for me to be able to read them all many books have been published recently*

(9) (a) **So many that I haven't been able to read them all books have been published recently*
(b) **Too many for me to be able to read them all books have been published recently*

Any account for result clauses that postulates a direct relation between *so* and the result clause has to face this problem. I will now try to show that it cannot be the case that the result clause is solely dependent on *so* or *too* respectively. I will argue that it depends on (i.e. identifies) the entire DP containing these elements. Assuming this we immediately get rid of the serialization problem concerning the ungrammaticality of (8-9). However, it cannot be denied that *so* and *too* do have an influence. They seem to induce the necessity of identifying an amount in terms of a result rather than in terms of another property (i.e. a relative clause). We therefore have to address the question as to how an element which is apparently not the head of the constituent under consideration has an influence on the nature of the identifier. The problem in general seems to be reminiscent of the phenomenon of pied piping. This becomes obvious if one looks at the constructions under discussion in comparison with wh-elements within pied-piped constituents:

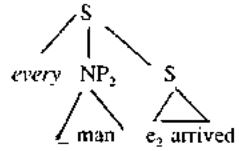
(10) (a) *[[so]$_{QP}$ [many [books]]$_{DP}$ that I haven't been able to read them all*
(b) *[with [whom]]$_{PP}$ did you leave*

The problem concerning both constructions is that an element that is more deeply embedded in a given constituent (*so* within DP in (10a) and the wh-constituent within PP in (10b)) behaves like it would be the head of that constituent. In pied-piping constructions the problem arises that the operator has to bind the variable (the gap of wh-movement) in order to get the correct interpretation. The wh-element must c-

[3] To be exact it might be possible that at S-structure they are adjacent to each other. One could postulate a PF-rule to be responsible for the 'anti-adjacency-requirement'. Here we would have an instance of an obligatory PF-rule.

command the variable. It is obvious that we are back to the problem of accessibility (of the feature of a non-head) as discussed in section III.6. There we have seen that two possible ways of analyzing this problem are possible. The first one (which is advocated in G&M for *so many books*) is to say that the quantifying element is moved out of its constituent at LF. Such a view is also found for the pied-piping problem. This sort of analysis is in a way reminiscent of Heim's (1982) approach for quantification in general. According to her any non-pronominal NP is adjoined to S (*NP-Prefixing* in her terms, *Quantifier Raising* in more recent terms). Then there is an additional rule of *Quantifier construal* which states that every quantifier is attached as a leftmost immediate constituent of S. For a sentence like (11) Heim assumes the following representation at LF (resulting in a tripartite structure that seems necessary for the interpretation we arrive at):

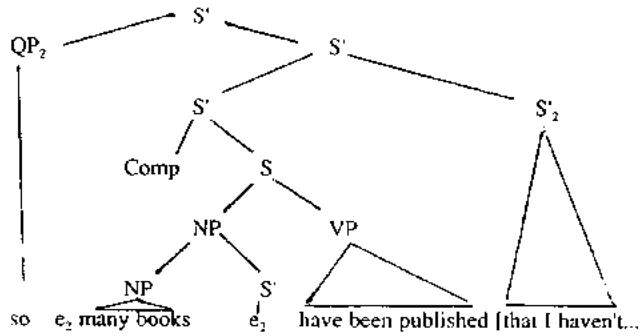
(11) (a) *Every man arrived*
(b)



(Heim, 1982: 133)

The same kind of analysis is advocated in G&M for *so*-NPs. The only difference is that *so* is assumed to attach to S' rather than S (i.e. CP rather than IP)[4]. Consider the LF-representation they assume for the sentence in (12a):

(12) (a) *So many books have been published that I haven't been able to read them all*
(b)



[4] As we will see this will account for the fact that result clauses have to attach higher than relative clauses.

In both approaches the whole constituent is first moved to its scope position at LF and then the purely quantifying element is further moved out of this phrase to its scope position. Now consider the relevant interpretation of quantifying elements:

(13) For every x [$_{RC}$x = a man] [$_{NS}$x arrived]

G&M's representation of *so NPs* in (12b) should then result in a similar tripartite structure:

(14) For so x [$_{RC}$ x = many books] [$_{NS}$ x have been published]

(14) does not seem to be a plausible candidate for the interpretation of *so NPs*. It seems to me that although some quantificational force is induced by means of *so* it is not true that this quantifier is restricted by the 'left-behind-NP'. The correct interpretation can just be obtained if the whole constituent is a unit (like it is also at S-structure). The LF-representation G&M assume to capture the syntactic facts does not give rise to an adequate input for interpretation.

Consider now the other solution for the pied-piping problem to be found in the literature. This is to say that the relevant feature (here the wh-feature) is somehow percolated further up. The entire PP inherits the wh-interpretation in case it is pied piped along with the wh-element. In this case the PP itself counts as the relevant operator binding the variable. In case of quantifying elements (and other elements that take scope over a given domain) such a percolation approach is found in van Riemsdijk & Williams (1981). They assume that scope is induced by virtue of percolation of the relevant feature[5]. I think that a percolation approach is preferable at least for pied-piping and *so- NP* constructions. Under this approach the constituenthood of the whole NP can be maintained, and as we have seen it is not just the syntactic behavior but also the interpretation of the relevant constituent that supports this (remained) constituency.

It has to be noticed that such an approach is also reminiscent of the compositional interpretation of nominals I am assuming in general. Remember that I assume that N is associated with an index X that counts as the index of the entire NP. Additionally D is associated with a second index x that is inherited by the entire DP. Once this

[5] Of course the problems to be addressed are the same. In both a movement and a percolation approach one needs to find the relevant restrictions for when and how far a certain feature or element can be percolated or moved respectively.

index appears index of NP itself is not accessible anymore (except N is stressed). In case of *so-NPs* the quantificational pronominal *so* can be assumed to be associated with an index of its own. Again this index is inherited by the entire constituent dominating it (i.e. the entire *so NP*). This assumption makes an important prediction, that creates a problem for G&M's analysis. If the noun is (contrastively) stressed it is possible to have a relative clause construed with a *so NP*. By means of stress on N X is accessible for Identification by the relative clause. In addition to the relative clause identifying X the entire constituent can be identified by a result clause:

(15) *Ich habe so viele MÄNNER gesehen, die gerne Bier trinken, daß mir ganz schlecht wurde.*
I have so many men seen, who like beer drinking that REFL totally sick got.

This example creates a problem for G&M (1984). They assume that the entire NP adjoins to S'. Thus the locality constraint imposed by the complement principle is violated in the example above. Now let me turn to the analysis of result clauses in the light of Identification.

I.2. Result Clauses as an instance of Identification.

Consider again the example (1) repeated here for convenience:

(16) (a) *Peter hat so viele Krügerl getrunken, daß er nicht mehr nach Hause kam.*
Peter has so many pints drunk, that he not anymore to home came.
'Peter has drunk so many pints that he couldn't go home anymore.'
(b) *Peter hat zu viele Krügerl getrunken, um noch nach Hause zu gehen.*
Peter has too many pints drunk, for still to home go
'Peter has drunk too many pints for him to be able to still go home.'

Obviously in (16) we do not find an individual picked out of a set of possible individuals by means of Identification as we found in the core cases of Identification. So we have to determine the properties of the identifyee and the identifier respectively. Descriptively we find the constituent *'so viele Krügerl'* as the identifyee and the result clause as the identifier. Crucially the identifyee contains *so*.[6] Notice that *so* is

[6] As we have already seen the identifyee can also contain *zu* ('too'). Here I will mainly concentrate on *so*. However, the analysis presented here should be equally available for the other possible (quantifying elements).

pronominal in nature - it can be used deictically, which is a crucial property of identifyees in general. Remember that it is a crucial assumption of the analysis suggested here that the identifyee has a representation at domain D - thus both the identifier and the identifyee should be associated with a discourse referent. What is the denotation in case of result clauses? Intuitively in this case the identifyee denotes a certain amount, whereas the identifier denotes the result of doing something (as denoted by the VP) with this amount. For example in the sentence above it is asserted that it was that much beer that Peter drank, that this amount sufficed to keep him from going home. This suggests that the result clause can somehow fix (i.e. identify) an amount. Let me briefly justify the claim that amounts can act as discourse referents. The first piece of evidence involves the already familiar test of pronominalization. We have argued that once we find the possibility to anaphorically refer to a previously mentioned constituent we can conclude that this constituent is associated with a discourse referent. Amounts are accessible for anaphoric resolution as can be observed in the following discourse:

(17) *Peter hat 9 Krügerl Bier getrunken. Maria hat auch so viel getrunken.*
Peter has 9 pints beer drunk. Mary has also so much drunk.

In the second sentence the pronominal element anaphorically refers to the amount of 9 pints beer. Notice that *so viel* cannot be interpreted as an elliptical measure phrase because it is not necessarily the case that Mary drank 9 pints. The sentence is still true if she drank the same amount of beer - not in pints but for example in small glasses.

Crucially there are two possible ways to denote an amount, one is *absolute* and one is *relative*. I will argue that comparatives are an instance of the former whereas in case of result clauses we find an instance of the latter. How can this be analysed? Remember that I make use of an extensional definition of properties, i.e. a property denotes the set of individuals satisfying this property. The denotation of an amount is then the set of those sets that have one particular property in common, namely the relevant amount. In the case above this would mean that *so viel* denotes the amount of 4,5 liters. That it is really this amount that is denoted can be seen on basis of the following example:

(18) *Peter hat 27 Artikel gelesen. Maria hat so viele Bücher geschrieben.*
Peter has 27 articles read. Mary has so many books written

The first sentence introduces an amount (of *articles*). *so viel* anaphorically refers to this fixed amount (i.e. *27 items*) and the noun adds another property the denoted set must

have, i.e. *being a book*. In other words the denotation of the amount above is the set of sets that have the property of containing 27 items (of something) in common. This is one possibility to denote amounts, namely an absolute amount, but there is also another possibility. It is a general property of properties that they can have certain *degrees*, i.e. they are associated with a *scale*. Individuals instantiating a certain property can vary as to which degree they have this property. Notice that adjectives (one instance of a syntactic category denoting a property) show this feature quite obviously: they can occur in *comparative* or *superlative* respectively. This means that properties are in general not (only) an absolute notion but they create scales according to which the property under consideration is evaluated. I will refer to the degree of a certain property as an *amount* of this property. If result clauses identify amounts then it is not surprising that it is not just absolute amounts (i.e. how many individuals are contained in set) but also degrees of a certain amount. Identification of a certain degree by means of a result clause means to fix a certain degree on the scale associated with the property under consideration. Along the lines of the representation we have been assuming for Identification so far I will assume the following representation for *so NPs* identified by a result clause:

(19) (a) *Peter hat so viele Krügerl getrunken daß ihm ganz schlecht wurde.*
Peter has so many pints drunk that REFL totally sick got

(b)
NP: X = set of individuals satisfying the property of being a pint
DP: x = many pints
so: Y = set of possible amounts (of many pints)
y = ?
$y \subseteq Y$
CP: y = that amount of pints that suffices to achieve the result that by drinking them, Peter got sick

This representation captures the relevant properties of Identification of *so NPs* we are interested in here. As usual NP denotes the set of individuals satisfying the property denoted by N. The determiner *many* denotes the discourse referent x which is a set of *many pints*. As I said *so* is pronominal in nature. Therefore it can be assumed that it is associated with two indices as well. I assume that Y corresponds to the set of possible amounts. The DP it is construed with is then reinterpreted as a property: it denotes a set of sets that contain a certain amount of pints that satisfies the property

of *being many* (pints).[7] *Many* is not an absolute notion. What counts as many is relativized to a given property. Therefore the set of sets of many pints could in principle contain any number of pints. I assume that this is an ordered set, the first member being the singleton set (containing just one pint of beer) the second set contains two members and so forth. This ordered set can be taken to correspond to a scale. This means that we can say that Y really denotes a scale of possible amounts. *so* is furthermore associated with a second index, which denotes the intended amount. In the representation above the denotation of *so* still contains an incomplete condition. The exact amount (drunk by Peter) is still not fixed. Again there is one more relevant information available, namely $y \subseteq Y$. The result clause identifies this incomplete condition. It picks out a certain set (on the scale). However this amount is still relativized to the drinking abilities of Peter. Identifying the denotation of y means to fix one particular amount on the scale that suffices in order to achieve the result that Peter cannot go anymore.

Notice that it is not only amounts in the sense of countable entities. Any property can be associated with a scale. It is therefore also possible to have *so* associated with a noun that contains an adjective:

(20) *Peter hat so ein gutes Bier entdeckt, daß er nicht mehr aufhören konnte zu trinken.*
Peter has so a good beer discovered, that he not anymore stop could to drink.
'Peter has discovered such a good beer, that he could not stop drinking anymore.'

What we find here is a set of individuals satisfying the property denoted by N. Then an intersection with the individuals satisfying the property denoted by the adjective is created. Subsequently *so* indicates that there is a certain degree, i.e. an amount of being good. If it is not clear from the present context how good the beer really is then it is identified by means of the result clause. The denotation of the identifyee is then the beer which is good enough to obtain the result denoted in the result clause.

It is now expected that this phenomenon is not only found for nouns modified by an adjective but also for 'bare nouns' since they also denote properties. We therefore predict that it is possible to have a result clause associated with a simple NP. This is

[7] Notice that this assumption is too simplified. Since I am not concerned with the representation of determiners in general I will abstract away from the problems induced by this notation. I will also not address the question as to how the presence of *so* induces a property reading again.

a correct prediction as exemplified below:[8]

(21) *Maria ist (so) eine Frau, daß ihr die Männer zu Füßen liegen.*
Mary is (so) a woman, that her the men to feet lie
'Mary is such a woman, that men admire her'

German allows for predicative NPs to be associated with a result clause. The presence of *so* is optional as expected under the present analysis. Any property denoting category can in principle be associated with a result clause, since the latter identifies a certain degree (i.e. an amount) of the property denoted by the identifyee. We might analyze the contribution of *so* along the following lines. It seems that the elements under consideration *so* and *zu* have the ability to make a property interpretation available. They can be associated with constituents that contain property denoting elements. Usually this property (i.e. predicative) interpretation induces a variable that is bound by the determiner. *so* seems to maintain this property interpretation. The interpretation of the nominal in (21) is such that the denoted individual is interpreted as to having the property of being a woman to such a degree that it obtains the result of men lying in front of her. We will return to some more properties associated with the scalar interpretation in the course of this chapter. For the moment the present remarks should suffice to argue that amounts are associated with a discourse referent.

Notice that all the amounts can be questioned. We might take the possibility to question a certain entity as a further test for being associated with a discourse referent:

(22) (a) *Wie viele Krügerl hat Peter getrunken?*
'How many pints did Peter drink?'
(b) *Wie gut ist das neue Lieblingsbier von Peter?*
'How good is the new favorite beer of Peter's?'
(c) *Was für eine Frau ist Maria?*
What for a woman is Mary
'What kind of woman is Mary?'

The question in (22a) can be answered by (17) the one in (22b) has (20) as a possible answer and (20) is an answer to the question in (22c).

Let me now turn to the justification of the claim that a result clause is associated with a discourse referent. One crucial evidence for this claim is the fact that nominals

[8] English does not allow for this option. I cannot give an account for why this should be so.

in general can refer to results. It is a well known property of nominalizations that they can be ambiguous between an event and a result reading. An example is given below:

(23) *Übersetzung*
translation

The derived noun can either denote the event of translating or to the result of this event, i.e. the actual book or paper translated. If nominals are always associated with a discourse referent then it follows that abstract entities like a result can be associated with as discourse referent.

We can also apply the already familiar test of anaphoric reference. Result clauses are possible antecedents for anaphoric resolution. This can be seen in the following discourse:

(24) S1: *Peter hat so viel Bier getrunken, [daß ihm schlecht wurde]$_i$.*
Peter has drunk so much beer that he got sick.
S2: *Maria hat auch so viel Bier getrunken.*
Mary has as well so much beer drunk.
(a) *aber [das]$_i$ ist Maria nicht passiert.*
but this is Mary not happened
(b) *aber [so weit]$_i$ kam es bei Maria nicht.*
but so far came it with Mary not

Notice that the possibility to anaphorically refer to the result clause by means of *so weit* (24b) indicates that it is really a certain point at a scale (a degree) which is fixed by the result clause.

Let me now summarize the core claim of the analysis the result clause as an instance of Identification. The identifyee establishes a set of individuals with a certain property. Properties in general establish the possibility of creating a scale. A certain point on this scale can be fixed by means of a result clause, i.e. the result clause identifies the amount. Since we are dealing with Identification on a scale it is a relative amount. It is evaluated on basis of some other property. In the example above it is evaluated on basis of the drinking ability of the subject. That it is really an amount relative the drinking abilities of the subject under consideration is should be clear from the example in (24). The absolute amount of beer is presumably the same in both cases (the second instance of *so much* being anaphorically related to the first one). However the drinking abilities of Mary seem to be higher. The very same (absolute) amount does in her case not suffice to end up in the result of getting sick. The result clause

fixes a point at a scale relative to the drinking abilities of the subject. The denoted amount is the 'sufficient' amount in order to arrive at the result denoted in the result clause. An amount can also be fixed w.r.t. some already established amount. This is so in case of comparatives. Consider the following example:

(25) *Peter hat so viele Krügerl getrunken, wie Maria.*
Peter has so many pints drunk, as Mary

Here the amount is fixed by comparing it to another amount (namely the amount of beer Mary drank). These remarks should suffice for establishing result clauses (and as we have just seen comparatives) as an instance of Identification. We will now turn to the predictions inherent to this analysis.

1.3. Identification is "optional".

It is a crucial property of Identification that it is not syntactically obligatory. The same is true for *so-XPs* and result (or comparative) clauses. If the denotation of *so-XP* is clear from the context then the identifier does not occur. This can be seen in the following examples:

(26) (a) *Peter hat so viele Krügerl getrunken (daß ihm schlecht wurde).*
Peter has so many pints drunk that REFL sick was
(b) *Peter hat zu viele Krügerl getrunken (um noch Auto fahren zu können)*
Peter has too many pints drunk for still car drive to can

Both sentences are wellformed (without an identifier) if there is a certain amount (of pints) which is salient in the discourse. Either it has been mentioned previously or a certain amount of pints is present in the discourse or the utterance is accompanied with an act of demonstration. Thus again we find the same pattern as in the core cases of Identification: the pronominal can be used anaphorically, cataphorically or deictically.

The following question answer pair (showing the same kind of phenomenon) is discussed in Williams (1974)

(27) Q: *Why wasn't the work finished?*
A: *Because so often no one was there*

"there is no Result clause, the so is undetermined. Its determination is actually the question itself." (Williams 1974: 196)

Identification is just necessary if the content of the pronoun is not already presupposed. Again we see that the relation between the pronominal element and the sentential element is not an obligatory licensing relation. If the content of the pronoun (i.e. the result) is not presupposed, or already established or salient in the discourse the occurrence of an identifier is again rendered obligatory.

1.4. Result Clauses and the Precedence Constraint on Identification.

If the relation between *so-XP* and the result clause is an instance of Identification we expect PCI to be obtained. This is a correct prediction as can be observed on basis of the following examples.

The identifier cannot be left-adjoined to the identifyee:

(28) (a) *Peter hat so viele Menschen, daß ihm ganz schlecht wurde gesehen.*
Peter has so many people that him very sick was seen
Peter has seen so many people that he became sick
(b) **Peter hat daß ihm ganz schlecht wurde, so viele Menschen gesehen.*
Peter has that him very sick was, so many people seen

The identifier cannot move to a position preceding the identifyee:

(29) (a) **Daß ihm ganz schlecht wurde hat Peter so viele Menschen gesehen*
that him very sick became was has Peter so many people seen
(b) *so viele Menschen, daß ihm ganz schlecht wurde hat Peter gestern gesehen.*
so many people that he got sick has Peter seen

A constituent containing the identifier cannot move to a position preceding the identifyee:

(30) (a) **gesehen, daß ihm ganz schlecht wurde, hat Peter so viele Menschen*
seen that he became sick has Peter so many people
(b) *so viele Menschen gesehen, daß ihm ganz schlecht wurde, hat Peter gestern*
so many people seen, that he became sick has Peter yesterday

Movement out of the identifier is impossible:

(31) **Wen hat Peter so viele Menschen gesehen daß er t anpöbelte*
Who has Peter so many people seen that he got mean to t

All the predictions of PCI are borne out in case of result clauses. Moreover, as expected the only counter-example involves Left Dislocation, i.e. the identifier can be left-adjoined to the CP containing the identifyee just in case the identifyee is topicalized:

(32) *Daß ihm ganz schlecht wurde, so viele Menschen hat Peter gesehen.*
That he became sick, so many people has Peter seen

Notice that the possibility to have result clauses in left dislocated position is further evidence for the claim that it is the entire constituent (containing *so*) that is identified by the result clause. Remember that the analysis for Left Dislocation proposed in chapter VII. suggests that a left dislocated constituent is licensed as a parasitic operator. As we have seen this derives the fact that Left Dislocation is only possible if the two adjacent constituents are coindexed. Assuming this generalization to be true we can draw the following conclusion for result clauses. Consider the representation of the sentences in (33):

(33) (a) *[Daß mir ganz schlecht wurde]$_i$ [[[so] viele] Biere]$_i$ hab ich getrunken*
that REFL totally sick was so many beers have I drunk
(b) *[Daß mir die Luft ausging]$_i$ [[[so] schnell]]$_i$ bin ich gelaufen*
that REFL the breath lost so fast am I run

so (although adjacent to the left dislocated result clause) is not itself coindexed with the trace of the topicalized constituent (structurally it is too deeply embedded). But the left dislocated constituent can only be licensed if it is coindexed with the trace of the topicalized constituent. We can interpret this as another indication that it is not just the pronominal element *so* that is identified but rather the entire constituent containing this *so*. Again it can be observed that if the entire constituent (i.e. the *so XP*) to be identified is more deeply embedded in another constituent then the sentence is not well-formed anymore:

(34) (a) *[Bücher von [so schlechten Autoren]$_i$] wurden kürzlich veröffentlicht, [daß ich sie gar nicht lesen wollte]$_i$.*
the books by so bad authors have been lately published that I didn't want to read them at all

(b) *[daß ich sie gar nicht lesen wollte]$_i$, [die Bücher von [so schlechten Autoren]$_i$] wurden kürzlich veröffentlicht.*
that I them not at all read wanted, the books of so bad authors have been lately published

Although it is possible to identify an identifyee that is more deeply embedded in another DP (34a)[9] Left Dislocation is not possible in this case. This follows from the present analysis since the left dislocated constituent cannot be licensed by being mapped to the trace in the matrix clause. In order to be licensed the trace and the left dislocated constituent would have to be coindexed. If the topicalized element only contains the coreferent element (instead of being the coreferent element) then coreference between the left dislocated element and the trace is blocked.

Notice moreover that in case of result clause Left Dislocation the pronominal is syntactically obligatory:

(35) (a) **[Daß mir ganz schlecht wurde] hab ich getrunken*
that REFL totally sick was have I drunk

(b) **[Daß mir die Luft ausging] bin ich gelaufen*
that REFL the breath left was I running

If there is no coreferent topicalized element then the result clause occupies SpecCP. It needs to be licensed by itself (not parasitically). Any topicalized element needs to be somehow inherently licensed within the clause it occurs in. A result clause like the one above is however not inherently licensed. It can just be licensed as an identifier.[10]

[9] Notice that this example apparently violates LCI. I will address this question in the next section.

[10] It is however interesting to notice that there are constituents (quite similar to result clauses) which are inherently licensed without identifying another element. In that case the presence of a prepositional complementizer seems obligatory:

(i) *Ich habe bis (daß) mir ganz schlecht wurde getrunken.*
I have until (that) REFL totally sick got drunk

(ii) *Ich habe getrunken bis (daß) mir ganz schlecht wurde.*
I have drunk until (that) REFL totally sick got

(iii) *Bis (daß) mir ganz schlecht wurde habe ich getrunken.*
Until (that) REFL totally sick got have I drunk

1.5. Result Clauses and the Locality Constraint on Identification

Let us turn to LCI. It has been noticed in the literature on extraposition that result clause extraposition seems less local than relative clause extraposition. This involves two paradigms. On the one hand result clauses seem to be adjoined higher as relative clauses. The common tests involve Condition C - effects:

(36) (a) *I told her that so many people attended last year's concert that I made Mary nervous*

(b) **I told her that many people attended last year's concert who made Mary nervous*

(G&M, 1984: 2)

Recall that I have argued that Condition C effects cannot be taken as a reliable test for the height of adjunction. Therefore I do not think that the examples above are significant for determining the height of adjunction.

On the other hand result clause extraposition can take place out of certain syntactic islands, but not so extraposition from NP as discussed in G&M:

(37) (a) *Plots by so many conspirators have been hatched that the government has jailed them.*

(b) **Plots by many conspirators have been hatched who the government has jailed*

Here we have a case where it is possible to relate an extraposed identifier with an identifyee that is 'more deeply embedded' than LCI would allow. However, we have already seen in case of extraposition out of NP that the domain can be extended by means of stress inducing focus. Under this light it is interesting to notice that *so* can also be used as a focus particle. It seems that by the same mechanism that is responsible for domain extension of extraposition out of NPs the domain can be extended in case of result clauses since there we have a focus particle involved.

This paradigm is similar to argument clauses since it is related to the verb directly and not to a mediating element (the correlate). In that case topicalization of the constituent itself is possible again. I do not have anything to say about the obligatoriness of a preposition.

Let us look at the locality properties more carefully. One prediction of the LCI is upward boundedness, which also holds for result clauses:

(38) *Daß Peter so viele Frauen gesehen hat, hat Maria nicht gewußt, daß er ganz nervös wurde.*
That Peter so many women seen has, has Mary not surprised, that he very nervous got.

The second Prediction of LCI is the Nesting Effect. Haider (1993) arguing for a strict ordering constraint for multiple extraposed elements mentions the following example that shows that the result clause is higher than all the other extraposed elements:

(39) *So insistierend hat sie jeden gefragt, der ihr begegnete, wenn sie gerade wieder mit Extraposition beschäftigt war ob dieser Satz grammatisch sei, daß die Leute stets verstummten.*
So insisting has she everybody asked, who she met, whenever she was again with extraposition busy, whether this sentence grammatical be, that people always kept quiet.

However, it is not surprising that the result clause here is the highest constituent since its correlate is also the highest element in the matrix. The word order in the example above is expected under the present analysis: it is always the case that the correlate in the matrix determines the adjunction site. We have to look at an example where the correlate for the result clause is deeper than an NP related to an extraposed identifier:

(40) (a) *weil der Mann so viele Bücher gelesen hat, daß er sich die Augen verdorben hat, der in dieser Bibliothek arbeitet.*
because the man so many books read has, that he himself the eyes ruined has, who in the library works

(b) **weil der Mann so viele Bücher gelesen hat, der in dieser Bibliothek arbeitet, daß er sich die Augen verdorben hat.*
because the man so many books read has, who in the library works, that he himself the eyes ruined has

Although as usual judgements are delicate with multiple extraposed constituents it seems to me that the nested ordering is again the preferred one. The data seem to indicate that domain extension is just possible constituent internally. As soon as the *main projection line* is reached there is no option anymore: the identifier must attach to the first maximal projection of the verbal projection line dominating the relevant constituent. I conclude that it is NOT a general property of result clauses that they are

adjoined higher than relative clauses. Notice that this is a problem for G&M's analysis. They predict that any relative clause invariantly has to precede the result clause whereas we predict the Nesting effect.[11]

1.5.1. Bound pronouns

The interpretation of pronouns as bound variables is often used as a test for the height of extraposition (cf. Haider 1993). I think that this test is really reliable. As I will show it supports the analysis suggested here. The pronoun within the result clause in (41) can receive a bound variable interpretation:

(41) *Jeder ist so schnell gelaufen, daß ihm ganz schlecht wurde*
Everyone is so fast run that he totally sick got
'Everyone was running so fast, that he got totally sick.'

The interpretation here is that for everybody it is the case that he ran that fast that he became sick. The bound variable interpretation can only arise if the relevant pronoun is in the c-command domain of the quantified expression. The present analysis predicts that in a sentence like (41) the bound pronoun interpretation can occur since the result clause has to be adjoined to the first maximal projection dominating *so-schnell*. The quantifier is hierarchically higher than the adjunction site for the extraposed element. The pronoun is within the c-command domain of the quantifier. The adjunction site of the result clause depends on the S-structure position of the correlate in the matrix. Notice that this is a problem for G&M's approach. If they would find a way that the quantifier can c-command the pronoun in the result clause they predict (under the assumption that QR targets the same position in any case) that any NP bearing quantificational force can bind a pronoun in a result clause resulting in the bound variable reading. This prediction is however not borne out as can be seen in the following example:

(42) **Wir sind so schnell jeder Frau entgegengelaufen, daß ihr ganz schlecht wurde*
We are so fast every woman towards-run, that herself totally sick became.

Here there is no bound variable interpretation available. The sentence cannot mean that for every woman it is the case that we ran so fast towards her, that she became sick.

[11] Again we can observe that English differs from German. As observed in Williams (1974) the position of the extraposed clause seems to be determined by scope properties of the correlate rather than the S-structure position.

The sentence is not even wellformed since the singular pronoun coindexed with the NP (*jede Frau*) is just possible if the bound pronoun reading is available. Otherwise a plural pronoun is necessary, and this is indeed possible:

(43) *Wir sind so schnell jeder Frau entgegengelaufen, daß ihnen/denen*[12] *ganz schlecht wurde.*
We are so fast every woman towards-run, that they got totally sick.

My analysis however makes the correct prediction: the S-structure position of *so* determines the adjunction site of the result clause. In the sentence in (43) the result clause is higher than the position of the quantifier at least at S-structure.
The same is also true if we change the relative order of *so-NP* and the quantifier:

(44) (a) **So schnell ist jeder gelaufen, daß ihm ganz schlecht wurde*
so fast is everyone run that he totally sick got
(b) *So schnell ist jeder gelaufen, daß ihnen/denen ganz schlecht wurde*
so fast is everyone run that them totally sick got

There is a clear contrast confirming our prediction: the bound variable reading is not available in (44a). Only the plural pronoun coreferent with the quantified phrase is wellformed. Let us now look at data involving the reverse test, namely a fronted object quantifier and a result clause with a pronoun:

(45) *Jeder Frau sind wir so schnell entgegengelaufen, daß ihr ganz schlecht wurde.*
Every woman$_{DAT}$ are we so fast towards-run, that her totally sick got.
'We were running so fast towards every woman that they got totally sick.'

Again my analysis makes the correct prediction. Here the pronoun in the result clause can receive the bound variable interpretation. *so-NP* determines the adjunction site. Here this adjunction site is lower than the position of the NP bearing quantificational force.[13] Notice also that the S-structure position of the quantified element is also relevant for scope assignment, i.e. *jeder Frau* is superior in scope to *so schnell*.

[12] Notice that the possibility for a d-pronoun is also expected because the antecedent does not count as a syntactic antecedent: it does not c-command the anaphoric pronoun. Therefore it is a relation of discourse antecedence (the only possibility for a d-pronoun to have an antecedent).

[13] I do not have anything to say about the fact that the S-structure position of the quantifier seems to be relevant for binding the variable.

1.6. Result Clauses and Presuppositions

We can now turn to the presuppositions induced by *so-NPs* and result clauses. We have seen already that *so* behaves similar to focus w.r.t. LCI. In previous chapters we have seen that stress on the determiner can extend the local domain for extraposition. *so NPs* always seem to allow for a less local domain. Moreover we have seen that stress on the determiner is necessarily associated with a certain presupposition. Consider now what happens in case of result clauses:

(46) *Ich habe SO viele Krügerl getrunken, daß mir schlecht wurde*
I have SO many pints drunk, that REFL sick got

For convenience let me repeat Bach's generalization that corresponds to the (forced) presupposition in case of stress on the determiner:

> "a restrictive relative clause presupposes the existence of entities of which the description given in the relative clause is not true. thus *the man that I saw* presupposes at least one man that I didn't see." (Bach, 1974: 272)

I have argued that this presupposition derives from the fact that the head NP is associated with two indices X and x. X corresponds to the set of individuals satisfying the property denoted by N. x (the actual discourse referent) has to be a subset of X therefore it follows that there is at least one man that does not have the property denoted in the relative clause. What could the parallel presupposition be in case we are dealing with a result clause? Remember that I analysed *so* as also being associated with two indices X,x. Here X corresponds to the set of possible amounts (i.e. the set of sets containing '*many pints*') that induce a scalar interpretation. Furthermore I have argued that the result clause identifies one amount at this scale. The denoted amount is the amount of beer that suffices to make Peter sick. Given that X corresponds to the set of possible amounts and x has to be a subset of x we expect the following presupposition in (47a) paraphrased as (47b):

(47) (a) Assertion: A_i NP R
Presupposition: There is some A_jNP (A_j is not A_i) such that ¬R is true of A_jNP, i.e. [A_jNP ¬R]

(b) There are *amounts of beer* of which the result given in the result clause is not true, i.e. there are *amounts of beer* that do not suffice that if I drink them I would become sick.

In (47) A corresponds to *amount* of NP and R refers to *result*. So the presupposition that arises is that there is at least some other possible amount that does not obtain the result denoted in the result clause. This presupposition is clearly found in a sentence like (46). It follows from the analysis of result clauses presented in the beginning of this chapter. *so* as a pronominal (focussing) element induces a set of possible amounts. The associated NP determines this set further. So we have a set of sets containing many beers. The result clause identifies the exact amount to be denoted, i.e. the denoted amount is the one that suffices for a certain result to obtain. Since the denoted amount is a subset of the set of possible amounts the presupposition arises that there are other amounts of beer that do not suffice that one becomes sick. Interestingly, in this case it is not any other amount that does not suffice to obtain the relevant result. As I have said the set of possible amount is ordered, i.e. it corresponds to a scale. The arising presupposition only concerns less big amounts. What the phenomenon under consideration is reminiscent of is *pragmatic scales* (in the sense of Fauconnier 1975). Later on we will see that the notion of *monotonicity* is related to the notion of pragmatic scales. This will be important for analyzing a generalization found in Williams (1974), namely that in English extraposition of result clauses can trigger Subject-Auxiliar Inversion.

1.7. Pragmatic Scales

Fauconnier (1975) assumes that the underlying structure of a sentence may not reflect its logical properties[14] directly and that such properties can result form inferences of a pragmatic nature. Secondly he claims that syntactic distributional properties may depend on these logical properties and therefore on the possibility of certain pragmatic inferences. It is important that the possibility for such inferences can be syntactically restricted.

His postulation of pragmatic scales is mainly based on the following observation. Superlatives (syntactic and pragmatic ones) can behave similarly to *any*, i.e. they can have quantificational force (in addition to a literal reading as a definite description). The following examples show this phenomenon:

[14] Of course this was stated at a time where the input for interpretation was supposed to be D-structure. However, I think that Fauconnier's main point is still valid within a different view on Grammar where some level 'after' S-structure (i.e. LF or domain D respectively) is taken to be the input for interpretation.

(48) (a) *Socrates can understand any argument*
(b) *Socrates can understand the most complex argument*
(49) (a) *Meno cannot understand any argument*
(b) *Meno cannot understand the simplest argument*

(48) shows that both *any* and the superlative *the most complex argument* can have a universal quantificational reading (i.e. *Socrates can understand every argument*). Whereas in (49), where the proposition is negated, both *any* and the superlative *the simplest argument* have an existential meaning (i.e. *it is not the case that there is an argument which Meno could possibly understand*). Moreover in existential environments both *any* and superlatives can occur if they are existentially interpreted however not if they are universally interpreted:

(50) (a) *There is not any/the faintest noise that he can stand*
(b) **There is any/the faintest noise that bothers him*
(c) **There is any/the faintest noise that my uncle can hear*
(d) *If there is any/the faintest noise that bothers you, please tell us*

The fact that a superlative can occur in *there* insertion sentences is rather surprising since normally existential contexts do not allow for definite NPs, but superlatives are definite. Fauconnier assumes that superlatives can function semantically to signal existential quantification.[15]

The question leading to the postulation of pragmatic scales is then why superlatives (can) behave like quantifiers? Consider again a sentence like the one below involving a superlative:

(51) *The faintest noise bothers my uncle*

The universal quantified reading can be explained if one assumes that if a certain noise

[15] Notice that Fauconnier shows that superlatives allow only for the literal (definite description) reading if they occur within syntactic islands:
(i) *Mary refuses to eat the most delicious food*
(ii) *Mary refuses John's suggestion to eat the most delicious food*
(iii) *Mrs. crabtree wouldn't let her daughter elope with the richest man*
(iv) *Mrs. Crabtree wouldn't let her daughter and the richest man elope*
(v) *John doubts that the most delicious wine is equal to his father's*
(vi) *John doubts that the most delicious wine and his father's are equally good.*
I take this fact to indicate that it is indeed necessary to represent possible presuppositions at some level of representation (which I take to be domain D) since they are sensitive to syntactic properties.

bothers somebody, then a louder one will too. Fauconnier thus assumes that the predicate *bother* is associated with a pragmatic scale, ranging from *faint* to *loud*, along the dimension *noise* as indicated below:

(52)

the loudest

x_1

x_2

the faintest (Fauconnier, 1975: 361)

Given the noises x_1 and x_2 such that x_1 is higher on the scale than x_2 (i.e. louder) it follows that the truth of *x_2 bothers y* will always entail the truth of *x_1 bothers y*. Moreover this assumption entails that the proposition will be true for all elements on the scale if it is true for the lowest element on the scale. The lowest element is most naturally identified by means of a superlative, in this case *the faintest noise*. According to Fauconnier such

> *"propositional schemata have the general form R(x,...). (The dots stand for possible additional free variables, such as y in x bothers y.) If x_2 is lower than x_1 on the scale associated with R(x,...), then R(x_2,...) implies R(x_1,...): thus, in particular if R holds for the lowest element on S, it holds for all elements of S."* (Fauconnier, 1975: 362)

Notice that the notion of pragmatic scales is reminiscent of the notion of *downward entailment*. It has been argued that downward entailing environments (i.e. if it is possible to infer from supersets to subsets) license the occurrence of polarity items (cf. Ladusaw 1980,1983). I will come back to this issue in the next section. What Fauconnier is able to predict with his analysis is that the logical negation of a proposition triggers the reversal of the corresponding scale, because

(53) if $R(x_2) \supset R(x_1)$ then $\neg R(x_1) \supset \neg R(x_2)$.

This means that the highest element on the scale associated with R will become the lowest on the scale associated with ¬R and the corresponding superlative will be used to convey the meaning of universal quantification. This is indeed the case as shown in the following example:

(54) *The loudest noise doesn't bother him*

For result clauses this means that it is just the case that the presupposed amounts which do not obtain the result denoted in the identifier are lower in the scale. Given that result clauses give rise to a scalar interpretation we expect that negation reverts the presupposition associated with (46) similar to Fauconnier's example involving superlatives. This is a correct prediction:

(55) *Ich habe nicht so viel Krügerl getrunken, daß mir schlecht wurde*
I have not so many pints drunk, that REFL sick got

The presupposition associated with (55) is given in (56a) paraphrased as (56b) which is the opposite as the one given in (47) above:

(56) (a) Assertion: $\neg A_i$ NP R
Presupposition: There is some A_jNP (A_j is not A_i) such that R is true of A_jNP, i.e. [A_jNP R]
(b) There are *amounts of beer* of which the result given in the result clause is true, i.e. there are *amounts of beer* that suffice that if I drink them I would become sick.

1.7.1. Subject Auxiliary Inversion with extraposed Result Clauses

Treating result clauses as an instance of fixing a certain amount on a given scale provides an interesting interpretation of a fact that was originally observed in Williams (1974). There it has been noted that

> *"no extraposition may take place from the complementizer position unless inversion occurs"* (Williams, 1974: 198)

The relevant examples discussed in Williams (1974) are given below:

(57) (a) *More often than I had anticipated John was not there*
(b) **More often than I had anticipated was John not there*
(58) (a) *Just enough to be suspicious, John was absent from work*
(b) *Too confused to accept the consequences, John committed suicide*
(59) (a) *Too often was John absent to finish his work*
(b) **Too often John was absent to finish his work*
(c) *More often was John absent than I had anticipated*
(d) **More often John was absent than I had anticipated*
(60) *So often was no one there that we went bankrupt*

In other words, extraposition seems to trigger Subject Auxiliary Inversion (henceforth SAI) in English. The relevant question we should address given this fact is: What triggers SAI in general? It would be a surprising and rather unmotivated result if it would be extraposition to trigger SAI. One reason to reject this conclusion is that in general extraposition does not trigger SAI. The second conclusion we would have to draw would be that SAI is a contextsensitive rule. But there is a more interesting generalization we can arrive at. SAI should rather follow from the nature of the topicalized element since this is usually the place where we should look for a trigger of SAI. Triggers for SAI in English are wh-movement (of an overt or a covert operator) (63). Topicalization in general does not trigger SAI except for some special constituents. Drubig (1992) identifies the following SAI-triggering elements: *no* and other negative quantified constituents, inherent negative adverbs of frequency (*seldom, rarely*) and expressions like *hardly, few, little, only*. The common property of the latter SAI-triggering elements is their monotonicity. Let us assume that this is the correct empirical generalization for the trigger of SAI: 'downward entailing' elements trigger SAI. The following examples exemplify this generalization:

(61) (a) *When has John kissed Mary*
(b) **When John has kissed Mary*
(c) *Has John kissed Mary?*
(d) **John has kissed Mary?*

(62) (a) **Yesterday has John kissed Mary*
(b) *Yesterday John has kissed Mary*

(63) (a) *No book has John ever read*
(b) **No book John has ever read*
(c) *Nowhere has John ever found a house*
(d) **Nowhere John has ever found a house*

(64) (a) *Only John has Mary ever kissed*
(b) **Only John Mary has ever kissed*

(65) (a) *Rarely did John go to the movies*
(b) **Rarely John went to the movies*

(66) (a) *Hardly anybody did John see*
(b) **Hardly anybody John saw*

It would be desirable if we could show that in the cases under consideration (i.e. examples (57-60) the same property (i.e. downward entailment) triggers SAI (no matter what explanation for this fact is chosen). The peculiar difference to other instances of SAI is however that downward entailment in these cases seems to depend on extraposition of the result clause. I think that the analysis I am proposing (making use of a scalar interpretation to derive the presuppositions) makes such an approach

available. If we want to stick to the generalization that it is the downward entailing property of the topicalized element that triggers SAI we have to claim that *so-XPs* (and other instances of identifyees construed with a result clause) are downward entailing if the identifier is in extraposed position but that they are not so if the identifier is directly adjoined to the identifyee. (Notice that if this generalization is on the right track we have again evidence that extraposition cannot be a PF-phenomenon, since it has a crucial interpretive consequence.)

Remember that we said that the identifyee denotes a set and the identifier fixes one entity/individual out of this set. The first thing we have to do is to check whether we really find a downward entailing environment. Consider:

(67) *Peter hat so viel Krügerl getrunken daß ihm schlecht wurde.*
Peter has drunk so many pints drunk that he got sick

The test for downward entailment is informally spoken to take a set and then take a subset of these objects in the same environment. If the first utterance entails the second the environment is downward entailing. Notice that this can be translated into the presupposition induced by pragmatic scales. There we found that once we talk about a certain point x_n at a scale then all other points (x_{n-x}) at the scale up to the established point x_n are implied. What does this mean for the example under consideration? First consider the interpretation of the sentence without the result clause *Peter has drunk so many pints*. Assume that *so many pints* refers to an amount of 9 pints. Drinking 9 pints entails that 8 pints were drunk, 7 glasses were drunk etc. So we can infer from a superset of amounts (9 pints) to a subset of amounts. So we can say that *drinking so many pints* is somehow a downward entailing environment. However consider the same clause in case the result clause is added: *Peter drank so many pints that he got sick.* What is asserted her is, that it was exactly the amount of beer that Peter drank that was sufficient for Peter to get sick. In this case it is not possible to downward entail anymore. If *so many pints* refers to nine pints and that was the amount to suffice that Peter got sick, the sentence does not entail that Peter drank 8 beers and got sick. Drinking less beer than the amount denoted does not end up in the result that is denoted in the result clause. This means that the result clause blocks the possibility for downward entailment. I think that this gives us an answer to the question we are interested here, namely why Subject Auxiliary Inversion is only licensed if the result clause is extraposed. The topicalized constituent is only downward entailing if the result clause is not part of the constituent. Once the result clause is added downward entailment (DE) is blocked.

(68) *So often* was no one there | *that we went bankrupt*
DE ¬DE

If the result clause is extraposed the identifyee (*so often*) is downward entailing. Therefore it licenses SAI. It is not before the result clause is uttered that downward entailment is blocked. If the result clause occurs in a position adjacent to the identifyee then SAI is not licensed because the whole constituent is not downward entailing. Notice that we have to make an additional assumption, which however we already justified. It seems that facts which have to do with discourse properties i.e. presuppositions (and now we have to treat entailments on par with presuppositions) are sensitive to S-structure properties.[16] Remember that I have argued in section II.3.2. that the Novelty-Familiarity condition is evaluated linearly in a step by step manner, i.e. we argued that interpretation is incremental. This has to be the case here too. Thus the difference between an extraposed and a non-extraposed result clause makes a difference as to whether the element in Spec CP (at the point when evaluation takes place) has its reference fixed. If the result clause is extraposed the exact amount of beer is not fixed yet. This creates a downward entailing environment. On the other hand if the result clause occurs in non-extraposed position the amount of beer is already fixed and the environment is not downward entailing. Notice that the entire sentence do not receive different interpretations, one being downward entailing and the other not. It is just the case that in case of extraposition a downward entailing environment is created for 'a little while'.

I think that the possibility to interpret this rather surprising fact that extraposition apparently triggers SAI provides us with an indirect support for our analysis. Identification and linear evaluation of the interpretation of discourse referents can explain this phenomenon without referring to construction specific statements.[17]

[16] Notice that another phenomenon which is sensitive to downward entailment is also sensitive to S-structure properties. Thus it is usually assumed that the licenser of a negative polarity item has to c-command the NPI at S-structure. Although problems with this generalization are easily found, I will not pursue this issue furthermore.

[17] It has to be noticed that there are still some problems inherent to this analysis. First, one should expect that polarity items are licensed in those environments.
(i) **So often was anybody there that we went bankrupt*
Polarity items are licensed in downward entailing environments otherwise. The present analysis would predict them to be licensed in case the result clause is extraposed. I have no explanation for this fact. Another problem for the present analysis is that apparently identification is necessary in order for SAI to be triggered:
(ii) **So often was no one there.*
Maybe this has to do with the fact that in case no identifier is added then the referent is fixed by

2. Conditionals

Another instance where I think that the relation of Identification postulated here is at play is conditionals.

2.1. Conditionals and Identification

As we have seen in chapter VIII. in case of conditional clauses the antecedent can occur in a position left-adjacent to the consequent. There we have argued (following Iatridou 1994) that if *then* occurs in adjacent (i.e. topicalized) position, then we are dealing with an instance of Left Dislocation. According to what we have said about Left Dislocation and its relation to extraposition we expect to find conditionals construed with a correlate in the matrix. This possibility is instantiated by the following example:

(69) (a) *[Wenn die Sonne scheint]$_i$, [dann]$_i$ wird Peter spazieren gehen*
If the sun shines, then will Peter walking go
(b) *Peter wird [dann]$_i$ spazieren gehen, [wenn die Sonne scheint]$_i$*
Peter will then walking go, if the sun shines.

(69) shows that the coreferent pronominal found in case of conditional Left Dislocation can also occur in the matrix clause if the conditional is in extraposed position. With the pattern in (69) we have a first indication that conditionals are really licensed as identifiers. In this section I will present some further arguments for this view. First we have to show that the conditional clause and the pronominal are associated with a discourse referent. We can make use of the already familiar test involving intersentential anaphoric relations. Consider in this light the following discourse:

(70) A: *Ach wäre das schön [wenn die Sonne scheinen würde.]$_i$*
'Gee, it would be wonderful if the sun would be shining'
B: *Ja, ja, [dann]$_i$ würde ich sogar mit dir spazieren gehen.*
'Indeed, then I would even go for a walk with you'

In (70B) the pronominal *dann* is anaphorically related to the 'if-clause' in the utterance of A. Therefore I will again assume that in example (70) the pronominal is associated

discourse.

with an index X and an index x to be identified. X denoted the set of possible antecedents. Since *dann* is associated with a temporal interpretation, I assume that this is the restriction on the possible antecedents. Informally we can say that *dann* refers to an eventuality at a given time. If it is used 'cataphorically' it is related to a set of possible events - which we might interpret as possibilities[18]. Out of some set of possibilities the actual possibility is identified by means of the clause in extraposed position. So we find the following representation:

(71) *Peter wird [dann]$_x$ spazieren gehen, [wenn die Sonne scheint]$_x$*
Peter will then walking go, if the sun shines.

Pron: X = set of possible eventualities
x = ?
x $\subseteq$ X
CP: x = if the sun is shining

I do not want to discuss this issue in a more elaborate way. Let me turn to some of the predictions.

2.2. Conditionals and 'optionality'

As we expect the presence of the identifier is syntactically optional. It is just necessary if the content of the pronominal is not clear from the context. On the other hand as expected the presence of the pronominal is not syntactically required - therefore it is also apparently optional:

(72) (a) *Peter wird dann gehen (wenn Maria kommt).*
Peter will then go (if Mary leaves)
(b) *Peter wird (dann) gehen, wenn Maria kommt*
Peter will (then) go, if Mary leaves

[18] Cf. also Asher (1994) for the claim that 'possibilities' can act as (abstract) discourse referents.

2.3. Conditionals and the Precedence Constraint

All facts predicted by PCI are also found in case of conditionals. The identifier cannot be left-adjoined to the identifyee:

(73) (a) *Peter wird dann, wenn Maria kommt gehen*
Peter will then, when Mary comes leave
(b) **Peter wird wenn Maria kommt dann gehen*
Peter will if Mary comes then go

The identifier cannot move to a position preceding the identifyee:[19]

(74) (a) *Wenn Maria kommt wird Peter gehen*
If Mary comes will Peter leave
(b) **Wenn Maria kommt wird Peter dann gehen.*
If Mary comes will Peter then leave
(c) *Dann, wenn Maria kommt, wird Peter gehen.*
Then, if Mary comes, will Peter leave

A constituent containing the identifier cannot move to a position preceding the identifyee:

(75) (a) *Gehen wenn Maria kommt wird Peter*
go if Mary comes will Peter
(b) **Gehen wenn Maria kommt wird Peter dann*
Go if Mary comes will Peter then
(c) *Dann gehen wenn Maria kommt wird Peter*
Then go if Mary comes will Peter

The last prediction is somehow irrelevant (i.e. it applies vacuously) since it is in general not possible to move out of the antecedent in conditionals:

(76) **Wer wird Peter (dann) gehen wenn t kommt*
Who will Peter (then) go if t comes

[19] Again the only counter-example is left-dislocation, as we have already seen.

2.4. Conditionals and the Locality Constraint

There are three main predictions that are captured with LCI: upward boundedness and the nesting requirement[20]. Moreover LCI is also relevant for data involving pronouns interpreted as bound variables since this is a test concerning the adjunction site of extraposed clauses.

First it can be observed that upward boundedness holds for extraposition of the antecedent clause in conditionals as well:

(77) (a) *daß Peter dann gehen wird, wenn Maria kommt war mir klar*
that Peter then go will, if Mary comes, was REFL clear
'That Peter would go, if Mary comes, was clear to me'

(b) **daß Peter dann gehen wird, war mir klar, wenn Maria kommt*
that Peter then go will, was REFL, clear, if Mary comes

My account for the nesting effects suggests that it is simply a byproduct of LCI (rather than a condition on chains, or an interpretive requirement). Since the S-structure position of the correlate in the matrix determines the adjunction site of the extraposed clause it follows that if the correlate of X precedes the correlate of Y then Y has to precede X if both X and Y are extraposed. This effect can also be observed with multiple extraposition involving conditionals:

(78) (a) *Der Mann würde dann Bier trinken, wenn das frische Faß aufgemacht wird, der ein Feinspitz ist.*
The man would then beer drink, if the new barrell opened will, who a gourmet is.

(b) **Der Mann würde dann Bier trinken, der ein Feinspitz ist, wenn das frische Faß aufgemacht wird.*
The man would then beer drink, who a gourmet is, if the new barrell opened is

[20] Notice that the behavior of extraposition w.r.t. a more deeply embedded antecedent is different here. It seems that whenever the pronominal correlate is embedded in an NP then the conditional reading is not available anymore. The interpretation induced is a purely temporal one. This is shown in the example below, where the sentential element related to *dann* cannot contain the (conditional) complementizer *wenn* but just the temporal complementizer *als*:

(i) **Die Aussage von dann wurde protokolliert, wenn Maria gekommen ist*
the assertion of then was recorded, when Mary come is

(ii) *Die Aussage von DANN wurde protokolliert, als Maria gekommen ist*
the assertion of THEN was recorded, as Mary come is

I have nothing to say about this fact.

(c) *Dann würde der Mann ein Bier trinken, der ein Feinspitz ist, wenn das frische Faß aufgemacht wurde.*
Then would the man a beer drink, who a gourmet is, if the new barrell opened is

(d) **Dann würde der Mann ein Bier trinken, wenn das frische Faß aufgemacht wurde, der ein Feinspitz ist.*
The man would then beer drink, if the new barrell opened will, who a gourmet is.

(79) (a) *Ich würde dann darüber nachdenken, ob ich kommen soll, wenn Maria nicht käme*
I would then thereover think, whether I come should, if Mary not came

(b) **Ich würde dann darüber nachdenken, wenn Maria nicht käme, ob ich kommen soll*
I would then thereover think, if Mary not came, whether I come should.

(c) *Darüber würde ich dann nachdenken, wenn Maria nicht käme, ob ich kommen soll*
Thereover would I then think, if Mary not came, whether I come should.

(d) **Darüber würde ich dann nachdenken, ob ich kommen soll, wenn Maria nicht käme*
Thereover would I then think, whether I come should, if Mary not came

The last two paradigms show that it is really the S-structure position of the correlated that determines the adjunction site of the identifier.

2.4.1. Bound Pronouns

As we have already seen one of the most reliable traditional tests for the adjunction site of extraposed clauses are quantifiers in the matrix binding a pronoun in the extraposed clause, resulting in the bound variable reading in case it c-commands the pronoun. Let us look at conditionals.

(80) *Jedem Schüler würde der Lehrer dann ein Buch geben, wenn er lesen könnte*
Every pupil would the teacher then a book give, if he read could

In this sentence the pronoun in the *if-clause* can have the bound variable reading, i.e. it can mean:

(81) For every x (x a student) (the teacher will give x a book if x has learned how to read)

The possibility for such an interpretation is predicted by our analysis since the adjunction site of the *if clause* is determined by *dann*.[21] It has to be adjoined in a position satisfying LCI. Therefore it is obviously lower than the quantifying element. However, if we change the S-structure ordering resulting in a configuration where the correlate of the conditional is higher than the quantifier a bound variable reading is not available:

(82) **Dann würde der Lehrer jedem Schüler ein Buch geben, wenn er lesen könnte.*
Then would the teacher every pupil a book give, if he read could

(82) is not wellformed, since the pronominal can just receive the group reading but not the bound variable interpretation. However a singular pronoun is not possible in this case. If we replace the singular by a plural pronoun the sentence becomes wellformed:

(83) *Dann würde der Lehrer jedem Schüler ein Buch geben, wenn sie lesen könnten*
Then would the teacher every pupil a book give if they read could

Since the group reading (thus the plural pronoun) is not necessarily structurally dependent on its antecedent i.e. it can pick *jedem Schüler* as a sort of discourse antecedent, there need not be a c-command relation between the quantifier and the pronominal element.

2.5. Conditionals and presuppositions

Now let us look at the presuppositions we expect if the relation is the same one as in the core cases. We said that the pronominal element introduces a set of possible eventualities. Again at the point of evaluation of the pronoun it contains an incomplete condition x = ?. The only information available is that $x \subseteq X$. According to what we have said so far, the following sentences should differ in the presupposition associated with them:

[21] Again with the proviso of section VIII.1.4.1. concerning the problem of why the S-structure position of the quantified element is relevant for binding the pronominal.

(84) (a) *Peter würde dann kommen, wenn Maria geht.*
Peter would then come, if Mary left
(b) *Peter würde kommen, wenn Maria ginge.*

The sentence with the pronominal element (84a) should have the following presupposition:

(85) There are eventualities/possibilities where it is not the case that Mary leaves

There seems to be a problem concerning this prediction. Witness the following contrast. If the particle *auch* ('as well') is added the presupposition indicated in (85) is forced. However adding *auch* is incompatible if the pronominal element is present:

(86) (a) *#Peter würde dann kommen, auch wenn Maria ginge.*
Peter will then come as well if Mary leave
(b) *Peter würde kommen, auch wenn Maria ginge.*
Peter would come, as well if Mary left

It seems to me that the actual presupposition induced by the pronominal is as follows. It is asserted in the consequence that Peter comes at a certain instance (out of the set of possibilities denoted by the pronominal. This pronominal is identified by the *if clause*. The presupposition that arises is that *only if Mary leaves will Peter come*. If however the proposition denoted in the *if-clause* is not true, then it is also not true that Peter will come. This presupposition does not occur if there is no pronominal element. Thus again we find a difference between the two sentences, i.e. there is no real optionality for the pronoun to occur, but again it has to do with discourse properties rather than purely syntactic constraints. Whenever we find the pronominal we get an interpretation which is equivalent to adding *nur* (only) (similar to what we found in case of Left Dislocation)

(87) (a) *Peter würde nur dann kommen, wenn Maria ginge*
Peter would only then come, if Mary left
(b) *Peter würde nur kommen, wenn Maria ginge*
Peter would only come, if Mary left

Notice that as expected *auch* can also not occur in the antecedent if the consequent contains *nur* (even if the pronominal is not present):

(88) *#Peter würde nur kommen, auch wenn Maria geht.*
Peter would only come, as well if Mary left

The presupposition we arrive at can be formulated like in (89a) or (89b) below:

(89) (a) out of the set of possible circumstances for Peter to come there is just one where it will really be the case that Peter comes.
(b) There is at least one other circumstance where it is not true that Peter will come.

(89b) is however quite similar to the presupposition Bach was describing with the only difference that it is not the property denoted in the identifying clause which does not hold, but it is the property denoted by the VP of the 'matrix' clause (i.e. the consequent). Again the presupposition is associated with the presence of the pronominal which is associated with two indices.

3. Some more instances of Identification.

Let me just briefly mention some other cases where I think that Identification is the relevant relation. Basically it involves all sorts of phenomena traditionally referred to as correlates. There are various possibilities involving all sorts of different relations to the matrix clause. Take the following instances as examples:

Adverbs of cause

(90) (a) *Peter ist (deswegen) gegangen, weil Maria gekommen ist*
Peter is (therefore) left, because Mary come is
'Peter has left, because Mary came.'
(b) *Weil Maria gegangen ist (deswegen) ist Peter gekommen*
Because Mary come is (therefore) is Peter come

Adverbs of place

(91) (a) *Peter hat Maria dort getroffen, wo er sie am wenigsten erwartet hätte*
Peter has Mary there met, where he here at least expected had
'Peter has met Mary, where he had expected her least.'
(b) *Wo er sie am wenigsten erwartet hat, dort hat Peter Maria getroffen*
Where he her at least expected hat, there has Peter Mary met

The properties discussed for the core cases as well as result clauses and conditionals are also found in these cases (i.e. the syntactic properties as well as the presuppositions). I would like to conclude that all cases involving a correlate in the matrix are best analysed as Identification at domain D.

APPENDIX I.
A discussion of previous analyses of extraposition

0. Introduction

Most of the basic questions concerning extraposition that have been addressed in the previous chapters are also addressed in previous analyses of extraposition. In this chapter I want to discuss some of the previous analyses and especially I want to compare them to the approach suggested here. We will see where and how they differ and in case my analysis makes different assumptions and predictions I will further elaborate the consequences and try to give an account for some apparent problems that arise through my analysis. Moreover, this section is also intended to justify some of the assumptions I have just taken from the literature, without discussing them (for example the claim that extraposition is not to be analyzed as an instance of move α).

It seems to me that there are two major problems with previous analyses: First the relevant questions are asked specifically concerning extraposition - resulting in more or less 'construction-specific' analyses. Secondly all the analyses have to stipulate a difference between left- and rightward relations. Both of these problems I think are overcome in the present analysis. Given the latter problem we are also confronted with one of the most discussed problems concerning extraposition. It has been noticed that there are empirical differences between left- and rightward relations (the most crucial one is that right-ward relations are at the same time more and less restricted than leftward-relations). Facing these empirical differences there are two possibilities to provide a solution. One could assume that there is a genuine primitive difference between left- and rightward relations. However, it would be a desirable result to derive the empirical differences from independently attested and motivated properties that distinguish the output of left- and rightward-relations. Notice that I have tried to give an analysis along the second line. However, one prerequisite for doing so is to assume that extraposed constituents are *'base-generated'* in the position they occur, without explicitly arguing for such a view. The next section is intended to support this view.

1. Movement or base-generation?

In giving a brief overview of the arguments for and against movement we can at the same time present an overview over some of the different analyses given for extraposition during the last decades.

1.1. Transformational Grammar

Within the system of transformational grammar it was assumed that deep structure is the relevant level for interpretation. To relate deep structure with the relevant surface structure several rules were proposed, all of them being language- and construction specific.

1.1.1. Ross 1967

One of the first most elaborate discussions of extraposition is found in Ross (1967). He states three different rules of extraposition that relate the sentences in (a) to the ones in (b):

(1) Extraposition from NP

X	$_{NP}$[NP - S]$_{NP}$ -	Y	
1	2	3 ----> OPT	
1	0	3+2	(Ross 1967: 1.10)

(2) (a) *A gun which I have cleaned went off*
(b) *A gun went off which I have cleaned* (Ross, 1967: 1.1-1.2)

(3) Extraposition[1]

X -	$_{NP}$[it - S]$_{NP}$ -	Y		
1	2	3	4	
1	2	Ø	4+3	(Ross, 1967: 4.126)

(4) *I claimed that Bob is a nut* (Ross, 1967: 4.127)

[1] This rule is followed by an obligatory deletion of *it* and restricted through the condition that extraposition may not apply vacuously.

(5) Extraposition of PP[2]

X -	[P NP]$_{NP}$ -	Y	
1	2	3 ----> OPT	
1	0	3+2	(Ross, 1967: 5.46)

(6) (a) *A review of this article came out yesterday*
(b) *A review came out yesterday of this article* (Ross, 1967: 5.45)

All these rules state that a specified constituent (S or PP) sister-adjoins to a variable to the right of the underlying position. The main goal in Ross (1967) is to find constraints on rules. This attempt can also be seen for the three rules of extraposition. His main claims concerning the restriction on extraposition rules are the following ones:

I. All Extraposition rules are upward bounded
II. All Extraposition rules are last cyclic
III. The three Extraposition rules cannot be collapsed

All these properties are still important for the discussion of extraposition in general. However concerning the question about the movement- or base-generation nature of extraposition the sort of answer given in the transformational framework is only important in order to see where and how the following discussion to be found in the literature departed from.

As it is clear from this brief summary of Ross' analysis, extraposition was supposed to be a transformation, thus a movement rule. However, there was no real possibility to assume extraposition to be a base-generated construction since as already mentioned it was assumed that a sentence which seems to be 'related' to another sentence is derived from the underlying configuration by the transformational component. Thus the question of base-generation vs. movement did not really arise.

1.1.2. Koster (1978)

One of the main steps to provide a basis for the discussion on movement vs. base-generation was the invention of traces (Chomsky 1977). Once traces are introduced the input for the interpretation of a sentence can be S-structure (as opposed to D-structure). Koster (1978) argues that the availability of traces undermines the traditional arguments

[2] This rule is stated as NP shift to generalize PP extraposition with Heavy NP shift

for movement rules. The question that arises then is whether the properties and conditions of movement rules differ from properties of rules of construal. Koster (1978) explicitly argues against a transformational approach to extraposition. He claims that it is possible for traces to be base-generated. (As we will see the discussion on movement vs. base-generation in the last few years is more a question on whether we find a trace or not).

At the time Koster (1978) argues for extraposition as a base-generated construction the main argument for extraposition as being derived by the transformational component (i.e. an instance of a movement rule) is that subjacency - a restriction assumed to hold for movement - seems to hold for extraposition. Various locality constraints on extraposition were derived from subjacency indicating that movement is involved. I will now present some of the relevant data. Chomsky (1975) for example gives the following data as evidence that extraposition obeys subjacency:

(7) Extraposition from NP
- (a) *The only one that I like of Tolstoi's novels is out of print*
- (b) *The only one of Tolstoi's novels that I like is out of print.*
- (c) **The only one of Tolstoi's novels is out of print that I like.*

(Chomsky 1975: 85)

In the examples above the sentence in (7c) violates subjacency as formulated in (8) (given the D-S as indicated in (9)):

(8) Subjacency Condition
No rule can relate X, Y in the structure
$...X...[_{\alpha}...[_{\beta}...Y...$ (or: $...Y...]_{\beta}...]_{\alpha}...X...$)
where α, β are bounding nodes.[3] (vanRiemsdijk&Williams 1986)

(9) $[_{NP}\ [_{NP}$ the only one that I like] of Tolstoi's novels] is out of print

A similar argument for extraposition obeying subjacency is given in Akmajian (1975) on basis of the fact that extraposition of PP prevents one reading of an underlying ambiguous NP:

(10) (a) *A review of a book by three authors appeared last year.*
(b) *A review of a book appeared last year by three authors.*

[3] The relevant bounding nodes assumed for English are S (IP) and NP.

The ambiguity of (10a) depends on where *by three authors* is attached, i.e. it can modify either *book* or *review*. However, example (10b) is just grammatical under the reading where *by three authors* modifies *review*. Subjacency excludes the other reading if one assumes the following derived structure:

(11) $[_{NP}$ a review of $[_{NP}$ a book t]] appeared by three authors.

In both types of examples (7) and (10) extraposition would involve movement crossing two cyclic boundaries (2 NPs). As Koster (1978) notices there are two problems with this approach. First, why can (7c) not be derived via (7b) i.e. why can the extraposed clause not first attach to the first NP and then move further to adjoin to its final position.[4] Second, referring to Guerón (1976) he shows that extraposition out of NPs is just possible under very limited conditions: only the complements of those NPs that are new information in a discourse may be extraposed, usually, these are indefinite NPs that are the focus of the sentence.[5] As for predicates, only very special ones qualify as contexts for extraposition. Extraposition from the subject NP, for instance, is only possible with *verbs of appearance*. Therefore (7c) is ruled out because no predicate of appearance is involved.

(12) **the novels t are out of print that I like most*

The sentence in (12) is not well-formed although there is just one NP boundary and the ill-formedness cannot be reduced to a violation of subjacency. Guerón furthermore argues that extraposed PPs or Ss are linked only to the focus of the sentence, and it is assumed by Koster that linking is possible without a movement transformation. In assuming this Koster can explain several data previously analyzed by means of subjacency.[6] Moreover, Koster argues that the *derived* PP position is a possible D-

[4] This problem is related to the well known distinction between right- and leftward relations, namely that extraposition cannot undergo successive cyclic movement (is upward bounded) but leftward movement can.

[5] Notice that this assumption is also a prerequisite for my analysis: Identification is assumed to be licensed in case the identifyee is a new discourse referent, obeying the Novelty Condition.

[6] On basis of the following examples he argues that this sort of linking is independently necessary, and not a construction specific rule.
(i) *A book appeared*
About relational grammar?
No about Raising again.
(ii) *This book is out of print*
**About relational grammar?*

structure position. Therefore he concludes that extraposition of PP is redundant - no necessity for assuming subjacency and thus no motivation for assuming a movement process. (We will discuss Koster's analysis for getting the locality effects as well as Guerón's Focus-based-linking approach later on.)

However, there is an even more 'full-fledged' base-generating approach available, i.e. we can even do without traces. If one adopts a strictly representational view of Grammar the question of movement vs. base-generation becomes a question on whether we do have to assume a trace or not. The first analysis that tries to make use of this 'fullfledged' base-generation analysis is the one given in Culicover & Rochemont (1990) (henceforth C&R). The locality constraint on extraposition, i.e. the subjacency and upward boundedness effects respectively are captured by the *complement principle*, which I will discuss at length in the course of this chapter. Let me just briefly give an overview of the development of this constraint.

1.2. Extraposition within the Principles and Parameters framework

1.2.1. Guerón 1980.

It was first suggested in Guerón (1978, 1980) that the relation between an extraposed constituent and the element it is linked to has to obey the following locality constraint (assumed to hold at LF):

(13) The complement principle:
The complement of X is a constituent governed by X

With this locality constraint (together with the Focus linking rule) Guerón tries to derive the properties previously captured by subjacency. However, Guerón still argues that extraposition is a movement rule (although the previous main argument namely that it obeys subjacency is not valid anymore). She claims that the following facts are at least indications the (PP)-extraposition is a movement rule.

(iii) *A review of a book appeared.*
By Bill again?
No, by Jill this time.
In these examples (involving question answer pairs with an effect best described by the linking rule - however, impossible to capture via subjacency) *by Bill* can only be linked to the entire NP. A discourse rule of Linking to Focus is argued to be independently motivated.

1. Move category and Structure preservingness. If the sole residue of transformation rules is *move category* (later on replaced by *move* α) then - so she claims - sentences like

(14) (a) *A man appeared with green eyes.*
(b) *John read a book over the summer by Chomsky*

will be automatically derived by the transformational component. She concludes that it would be redundant to generate them in the base as well.[7] Obviously this line of argumentation implies that extraposition is a non-structure preserving rule. Structure preserving rules can just move an element to a position where it could also be base-generated. (In this case the assumption of movement would be undermined - as Koster (1993) notices). Guerón is aware of this implication and she goes even further to state that the non-structure-preserving nature of the rule argues for its existence. Before we will see that even within the framework Guerón is assuming the argument presented above is not valid, I would like to discuss briefly whether it is still a valid argument within the Principles and Parameters framework.

The notion of structure preservingness was introduced in Emonds (1976) and is defined as follows:

(15) Structure Preserving Constraint:
Major grammatical Transformations are either root or structure-preserving operations.
(16) Structure Preserving Transformation:[8]
A transformation that introduces or substitutes a constituent C into a position in a phrase marker held by a node C is called structure-preserving.

[7] It is interesting to notice at this point that C&R argue against the assumption of a trace left by extraposition on exactly the same line of reasoning: If linking is done anyway then it is not necessary to assume a trace as well, since the principle of full interpretation is fulfilled without postulating a trace. We will return to this issue later on.

[8] For completeness see the following relevant definitions given in Emonds (1976):
Root Sentence:
A root S is an S that is not dominated by a node other than S
Root Transformation:
A transformation that moves, copies or inserts a node C into a position in which C is immediately dominated by a root S in derived structure is a root transformation.
Local Transformation:
A Transformation that effects only an input sequence of a single non-phrase node C and of one adjacent constituent C' that is specified without a variable, such that the operation is not subject to any condition exterior to C and C', is called a 'local transformation'.

This amounts to saying that *structure preserving transformations* are transformations that derive structures that would be in principle allowed to be derived in the base - i.e. they neither build new structure nor do they change structure. Obviously it is crucial to determine what counts as a possible base-structure. Guerón (1980) claims that it is not a possibility (in Dutch) to have PPs following the verb in final position in the base and therefore she concludes that extraposition is not structure preserving. This is however a conclusion which is not valid. Since adjunction is not only a configuration which can be derived via movement but base-generated adjunction is also a possibility Guerón's necessary presupposition is undermined. I.e. it is not clear why extraposed PPs even in Dutch should not be base-generated in the position they occur in. (Recall that Koster 1978 explicitly argues that they can.) It is true that extraposition cannot be substitution (i.e. structure-preserving in the pure sense) because then one would have to assume a specifier position to the right. Moreover it would have to be assumed that there is a projection (with such a right-ward specifier) between every projection (remember that one can show that there are several different positions available for extraposition). But this fact as well as the fact that it is not possible for any constituent to occur in the position following the verb does not necessarily imply that it has to be a derived structure. We just have to determine whether an element is allowed (i.e. licensed) in right-adjoined position or not - under the base-generation analysis it must already be licensed there at D-structure under a movement-approach it must be licensed there at S-structure. The difference between the two approaches can just be detected on basis of finding a trace or not and on basis of finding some real constraints holding for movement. As we will see below, the assumption that extraposition leaves a trace creates several problems (no matter whether subjacency or the complement principle is assumed to derive the locality constraint).

I therefore conclude that the first argument for the movement-approach is not valid anymore in the framework we are dealing with now. In addition there is a second empirical problem for this argument (which also constitutes a problem within the transformational view).

2. *Split Antecedents*. Guerón argues that the impossibility for extraposed PPs and relative clauses to have a split antecedent is an argument for the movement approach:

(17) (a) **A man met a woman yesterday from two different regions of India*
(b) **A man met a woman yesterday who were similar.*

There are two important points to criticize here. The first point is mentioned in Guerón (1980) in a footnote (citing Chomsky p.c. for this conclusion). It is

"the fact that Comparative Extraposition and Consecutive Extraposition do allow split antecedents, as in

(i) *So many boys were talking to so many girls that the room was very noisy*

[which] suggests that these last structures are basederived."

(Guerón, p.648,Fn.12).

However, this argumentation undermines Guerón's first (and presumably main) argument, i.e. that the possibility to base-generate Ss (and PPs) in right-adjoined position would be redundant. If however the base allows for Ss to be right-adjoined it should not play a role what sort of S we are dealing with. And if this sort of configuration is already possible in the base, then extraposition would not be non-structure-preserving - thus its non-structure-preservingness cannot argue for its existence.

The second point (which is related) is the following: it has been observed in Perlmutter & Ross (1970) that even extraposed relatives (and not just comparative and consecutive extraposition) can have split antecedents:

(18) *A man entered the room and a woman went out who were quite similar.*

Given the grammaticality of this example along the same line of reasoning as in Guerón's footnote we can say that we have to deal with a basederived construction. If so, then again the first argument of Guerón completely fails, even for relative clause extraposition and the same is true for the second argument. Thus we are left with a last (theory-internal) argument for the movement approach mentioned in Guerón.

3. The Name Constraint. In order to derive the impossibility to extrapose out of certain NPs Guerón's analysis crucially hinges on the assumption that a trace is left in the relevant cases. She states that a complete referring expression cannot contain a variable that is not bound within it (i.e. the *Name Constraint*). Since extraposition would be an instance of A'-movement its trace would be a variable and movement out of a complete referring expression is prohibited. (Thus with her analysis Guerón tries to capture the fact that extraposition is not possible for non-restrictive constituents.)[9] Of course the argumentation here is theory-internal and crucially hinges on the empirical and theoretical adequacy of the Name Constraint. However, there is counter-evidence

[9] This follows since it is always the case that a non-restrictive relative clause is related to a full referring expression, i.e. it is not an inherent feature of a relative clause to be restrictive or non-restrictive, rather it is a matter of whether the head NP is a full referring expression. If it is then the relative clause can obviously not be interpreted as restrictive.

against its theoretical and against its empirical adequacy.

Although Guerón does not give an exact definition of Names and the Name Constraint, the following theoretical question still arises. Traces of A'-movement are defined as variables and variables are considered to behave on a par with R-expressions. This assumption however could cause a problem for the Name Constraint, namely if 'being a trace' would be equated with 'containing a trace'. Especially under a view where traces bear categorical information (which is necessary under the assumption that there is no relevant distinction for binding theory between overt and covert NPs) then the representation of the variable would be $[_{NP}\ t]$, which looks like an NP containing a trace - thus contradicting the Name Constraint.[10]

An empirical problem for the Name Constraint is the fact that it is possible to extrapose out of definite fully referring NPs. Especially the examples we discussed arguing that a definite NP introduces a discourse referent are relevant. In that case the definite DP refers to a unique individual by virtue of the identifier (as we have argued). Thus I take it that the Name Constraint has empirical as well as theoretical problems. An additional theoretical problem has to do with the fact that Baltin (1987) argues that the assumption of a trace causes problems for Guerón's analysis. If the trace has to be deleted (which Baltin argues to be a necessary conclusion), then the Name Constraint could not really exclude the data Guerón wants to capture with it.

Of course if one does not make use of a movement analysis of extraposition an adequate analysis for the facts captured by Guerón's Name Constraint must be captured by some explanation which provides us with a prohibition on extraposed elements to relate to elements having a certain property. As we have seen the analysis suggested here crucially predicts such a difference. A complete referring expression can never induce the need for Identification. In these contexts extraposed constituents can never be licensed as identifiers. However, although our approach can capture the data the Name Constraint is intended to capture it is not the case that we capture exactly the same empirical domain. Notice that both of the problems for the Name Constraint discussed above do not create a problem for the Identification analysis. We can dispense with Guerón's Name Constraint. I therefore conclude that there are no

[10] As mentioned before, it is not clear to me what really counts as a Name (the definition being that a Name is a full referring expression). One could do away with the problem just mentioned if one would claim that Names are only a subset of R-expressions. It seems to me that the definition Guerón gives can explain the behavior of real variables, but not of variables that are left behind for example by topicalization.

arguments left for the assumption that extraposition has to involve movement. Now let us turn to the positive evidence that extraposition is in fact basegeneration.

1.2.2. Baltin (1981-1987)

Paradoxically as it may seem evidence for the base-generation approach can be drawn from Baltin (1987) arguing explicitly for extraposition as an instance of movement obeying subjacency. Two arguments are relevant here: one is his argument against Guerón and the problems her complement principle has to face if one assumes traces. The other is drawn from a critique given in C&R concerning Baltin's analysis. Let me just briefly summarize the main points in Baltin's analysis. He tries to reduce the more restrictive property of extraposition (as opposed to leftward movement) to a specific definition of subjacency:

(19) *Generalized Subjacency*
In the configuration $X.....[_A.....[_B.....Y...]...]...X'$,
(a) X and Y cannot be related where A and B = one of NP,PP, and either or both of S and S'
(b) X' and Y cannot be related where A and B are both maximal projections

On the one hand under this definition rightward movement is more restricted than leftward movement as the nodes counting for subjacency are assumed to be any maximal projection dominating the base-position whereas for leftward movement specific bounding nodes are mentioned. On the other hand Baltin suggests that the less restrictive behavior of extraposition can be reduced to the fact that extraposed constituents have the character of adjuncts. Making use of the assumption in Lasnik & Saito (1984) that movement leaves traces just in case they are required by independent principle of Grammar, Baltin (1987) suggests that traces of extraposed phrases need not be visible to satisfy any principle. Therefore the ECP is not relevant for extraposition. To conclude: strict boundedness is reduced to (a specific formulation of) the subjacency requirement, the possibility for rightward movement out of subject NPs is reduced to the adjunct character of the extraposed phrase.

First let me mention that both assumptions can be rejected on empirical reasons. For the first assumption concerning Generalized Subjacency it has to be noted that upward boundedness is not obeyed if the extraposed phrase moves out of an infinitival complement:

(20) (a) *Peter hat [dem Mann den Computer zu bringen, [den er nicht mehr braucht]] versprochen*
P. has the man DAT the computer (ACC) to bring [which he not anymore needs] promised
(b) *Peter hat [dem Mann den Computer zu bringen] versprochen, den er nicht mehr braucht*

(21) (a) *Peter hat dem Mann zu sagen versprochen, wann das Fest stattfindet.*
P. has the man DAT to say promised, when the party takes place
(b) *Peter hat dem Mann zu sagen, wann das Fest stattfindet verprochen.*
Peter has the man to say when the Party takes-place promised

Either an argument CP (21) or a relative clause (20) can be adjoined to the infinitival clause or to the matrix CP. Furthermore the left- right asymmetry has a stipulated character (which Baltin also admits but cannot dispense with to arrive at an empirically adequate theory of extraposition). It has to be assumed that there is a genuine primitive difference between left- and rightward relations, a result we want to avoid from the beginning. Moreover he can not explain why extraposition cannot undergo successive cyclic movement (like it is possible for rightward movement).

The second problem is that not all extraposed constituents are adjuncts (as also mentioned in a footnote in C&R (1990)) and it is not plausible why the projection principle should not demand the trace of the extraposed element. C&R (1990) discuss the problem of assuming the deletion of the trace in general. They consider it as a violation of the Principle of Full Interpretation. In their view an extraposed constituent is neither an argument of the head of the projection to which it is adjoined, nor is it a predicate nor an operator. It can therefore only receive an interpretation via the trace (if a movement account is pursued).

Let me now summarize the facts leading to the conclusion that extraposition does not involve movement. First I have shown that some version of the complement principle (although formulated differently in the account suggested here, namely the Locality Constraint on Identification) can explain the locality conditions on extraposed elements. If so then subjacency as a constraint holding for extraposition can be dispensed with. Secondly I have argued above that a representational view of Grammar suggests that the question of whether we are dealing with movement or with base-generation amounts to asking whether we find a trace or not. Thirdly Baltin explicitly argues against positing a trace for two reasons: First, it is a necessary assumption within his analysis in order to derive the less restrictive behavior of extraposition. Secondly Baltin also argues against positing traces in an approach which makes use

of some version of the complement principle (I will present his critique below). Therefore we can conclude that extraposition does not involve movement but is fully base-generated.

Baltin's argument against postulating traces is as follows. The data discussed in Guerón indicate that she has to assume that only overt constituents can count as complements for the complement principle. This is so since the trace of the extraposed element would be governed by constituent left behind (the head which has to govern its complement), and the unacceptable sentences she wants to exclude by the complement principle would all be predicted to be grammatical. But this assumption - as Baltin convincingly argues - leads Guerón to an inconsistency since traces from leftward movement must be allowed to be visible for the complement principle in order to get a correct result. Otherwise a moved complement would not be governed by its head anymore. Guerón also has to make a distinction between left- and rightward relations and is therefore not superior to assuming two different versions of subjacency - one which is obeyed by leftward relations and one which is obeyed by rightward relations. However, if leftward relations (as opposed to rightward relations) need to leave a trace (they are movement relations in most cases - a matter we will return to) but rightward relations do not leave traces then the problem mentioned in Baltin does not arise. In arguing against the complement principle Baltin argues against the movement nature of extraposition. Finally let us look more closer at the argument against the movement nature of extraposition discussed in C&R.

1.2.3. Culicover & Rochemont (1990)

The first argument against Baltin's approach has already been mentioned, namely that it leads to a contradiction to assume that the trace of extraposition (considered to be an adjunct) is deleted since this would violate the Principle of Full Interpretation (PFI). The way out C&R are proposing is to assume that

> *"an extraposed phrase could receive an interpretation without binding a trace within the NP with whose head it is construed. Then the PFI could be satisfied without any need for postulating a trace within NP and the lack of ECP effects with extraposition could be properly accommodated without jeopardizing the ECP account of the restricted interpretation of dislocated adjuncts."* (C&R, 1990: 26)

At this point the complement principle suggested in Guerón comes into play again: it is supposed to be responsible for the extraposed element satisfying PFI. For doing so

they claim that if the extraposed element is in a government relation with the head it is construed with, PFI is satisfied. The extraposed element can be interpreted properly. By Occam's razor (in addition to the above discussed problems with the assumption of traces) C&R conclude that no trace is needed and thus extraposition is better analyzed as an instance of a base-generated construction and the complement principle is preferable over the subjacency account.

This conclusion seems to be supported by the following typological observation cited from Downing (1978):

> *"In other languages, as we have seen, co-relative clauses may be structurally identical with adverbial clauses rather than being associated with ad-relative clauses in the same language. This situation suggests again that co-relative clauses should not be looked upon in general as merely relocated variants of embedded clauses."*[11]
>
> (Downing 1978: 410)

Moreover for 'extraposed' relative clauses in Walbiri he states that:

> *"There is no syntactic evidence that Rel NP forms a constituent with the modified NP at any level."* (Downing 1978: 406)

Judging from this generalizations the claim that extraposition is not derived by an instance of move α is supported also on basis of cross-linguistic evidence.

1.3. Movement vs. basegeneration and left- vs. rightward relations

Having argued for the base-generation approach towards extraposition we have to address the question whether the sort of arguments presented so far would also imply that (certain) leftward relations are base-generated rather than derived via movement. Here I simply want to refer to the analysis presented here where the output referred to as extraposition is a result of the Precedence Constraint interacting with the Locality Constraint on Identification. The basic question is always whether and how an element is licensed in the position it occurs in. If for example an element occurs in SpecCP it is not local enough to be licensed via Identification thus it can just be licensed by a trace (which is in most cases also necessary for satisfying the projection principle) in

[11] ad-relative clause refers to relative clauses which occur adjacent to the head NP, whereas co-relative clauses refer to 'dislocated' (i.e. extraposed) relative clauses.

order to be mapped from its A'-position into an A-position. Assume that a potential identifier is adjoined to a position satisfying LCI but is linearized to the left rather than the right of the main projection line. Then this identifier would be in the very same hierarchical configuration. This output should in principle be possible (and it is possible in case there is no identifyee, as we have seen in chapter VI.). However, this output violates the PCI. So I take it that for all the relevant examples discussed here there would be no possible leftward-relation to be established in order to license the left-adjoined element as base-generated. Notice that C&R cannot derive this. Under their approach it should not make any difference whether the 'complement' is adjoined to the left or to the right, the only constraint being that it is in a government relation to its head.

Let me briefly mention the only other attempt to capture the difference between left- and rightward relations, i.e. the only account that tries to derive the prohibition against *leftward-movement* in case *rightward-movement* is allowed. Such an account is given in Guerón (1980). She claims that the obligatory rightness does not have to be stipulated. In principle move α would allow for the relevant elements to be moved leftwards. Guerón assumes that independent constraints rule out leftward movement in the appropriate cases. However, she does not really give a satisfying account for those independent principles. She claims that an example like:

(22) $[_{S_0}\ |_{PP_1}$ with green eyes$||_{S_0}[_{NP}$ a man $|_{PP_1}e]]$ appeared (Guerón, 1980: 639)

is filtered out by an interpretive rule having the following effect:

(23) In the structure $X^n\ |_{S_0}$ NP...|, where n indicates a number of bars, X^n is the Topic of S^0.

Given this rule of interpretation she concludes that on the level of complete semantic interpretation, (22) will then be ruled out because *with green eyes* is not a suitable Topic for the S *a man appeared*. Notice that my analysis makes a different prediction. In fact it does not rule out (22). PCI cannot apply because PP and NP are not coindexed in the same way as other identifiers are, but PCI is reduced to the Novelty Condition and therefore it can only apply if the two constituents are really coindexed. In the case above the PP does not bear the same index as the correlated NP. So there is no previous occurrence of this very index and Novelty is satisfied. This also means that the topicalized PP is licensed differently - it is not an instance of Identification in the same way as the core cases discussed here are. Moreover, here we presumably are dealing with an instance of movement, i.e. the PP is moved to Spec CP and therefore

inherently licensed because it is in a chain-configuration with the trace in base-position. So we expect (22) to be well-formed. It seems to me that in German this is really what we find. If it is in principle allowed to move out of an NP then a configuration like (22) is well-formed:

(24) (a) *?Mit grünen Augen habe ich einen Mann gesehen.*
with green eyes have I a man looked-for.
(b) **Der grüne Augen hatte, habe ich einen Mann gesehen*
who green eyes had, have I a man seen

As we expect, there is a striking contrast between (24a) and (24b) predicted by the present analysis. Topicalization is only ruled out (by PCI) if the topicalized element and its correlate are coindexed.[12]

Moreover, notice that Guerón's approach only captures a subset of the empirical domain PCI captures, i.e. it cannot explain why it is not possible to topicalize an otherwise appropriate topic as soon as the identifier is topicalized along with this constituent but the related element is left behind in the matrix. I take it that PCI, that is reduced to the Novelty Condition is empirically and theoretically superior to Guerón's account, which is however, the only attempt I am aware of to address this problem - all other analyses have to stipulate the obligatory nature of rightness in the cases we are interested at.

To sum up leftward relations which are assumed to be created via movement can be divided into: A'-relations, A-relations and X°-relations. For all of those relations postulating a trace is unavoidable simply given the projection principle. In addition for A'-relations postulating a trace is necessary in order to make sure that case-assignment can take place. For A-relations the trace is in the position where theta-assignment takes place - thus the theta-criterion is just satisfied if a trace is assumed. And for X°-movement the trace is necessary since otherwise we would get a result where a projection would not have a head. On the other hand for rightward relations it is the case (at least in German) that it is just possible for elements which do not need case to occur in extraposed position. All of the relevant examples (as I have argued also apparently complement CPs) are licensed as identifiers rather than being directly theta-marked. Thus the main reasons for postulating a trace within a relation are not given in rightward-relations.

[12] It was pointed out to me by Edwin Williams (p.c.) that the sentence in (22) (judged ill-formed by Guerón) is well-formed in case the PP is interpreted as non-restrictive.

2. Locality Restrictions on Extraposition

As we have already seen above extraposition shows several locality constraints. One of them has already been discussed briefly, i.e. the subjacency effect (captured via the complement principle in Guerón (1980), Guerón & May (1984) as well as C&R or by the LCI suggested here). Since we have argued against extraposition as a base-generated configuration I will not go into the subjacency accounts any further. Let me just mention that one advantage of some version of the complement principle (i.e. a wells-formedness condition deriving the correct result rather than a constraint on movement) is that it can capture both, the 'subjacency effects' and the apparent failure of rightward movement to undergo successive cyclic movement. Let me briefly discuss this locality constraint and its predictions.

2.1. Upward boundedness

Upward boundedness has first been introduced in Ross (1967) as cited above. This property has been reduced by a constraint involving command. Later it has been tried to be derived via subjacency. However, there is one more problem for an account like that. Upward boundedness does not hold for leftward movement, i.e. leftward movement can take place successive cyclically via SpecCP (acting as some sort of an escape hatch). This option does not seem to be available for rightward movement although there is no a priori reason as to why SpecCP should not be available for rightward movement - as in terms of hierarchy left and right should not make any difference.[13] There are two different ways out of the problem to be found in the analyses involving movement.

Ross tries to explain this fact by claiming that extraposition is a last-cyclic rule. There has been a big discussion in the seventies on the cyclicity or last cyclic nature of extraposition respectively (cf. section VI.1.4.2. for a discussion). The main result of this discussion was that extraposition is a cyclic rule (contrary to Ross' claim). As we have already seen most of the arguments can be taken as arguments against extraposition as PF-phenomenon. If extraposition is a cyclic rule then the possibility for moving through intermediate SpecCPs should not be in principle excluded.

[13] See section 3. for a discussion of Müller (1993) who claims that there is a genuine difference between left and right, however, without motivating this approach.

Baltin (1987) is aware of this problem (especially given the argument in Guerón 1980 for extraposition as a syntactic rule as apposed to a stylistic rule) and argues for the following way out (already suggested in Baltin 1982). He posits the following notion of proper government for the ECP:

(25) A properly governs B iff A minimally c-commands B and
  a) A is the antecedent of B, or
  b) B is contained within a maximal projection of A, and A selects B.

Assuming this definition successive cyclic movement is blocked by the ECP since the intermediate trace, not being c-commanded by a category that selects it, will not be properly governed. It is crucial for this approach that extraposition is a movement rule. Besides the genuine problem with the assumption that extraposition is a movement rule there is another problem with the approach suggested in Baltin. It has to be assumed that intermediate traces of leftward movement behave differently in order to get the result that leftward movement is not upward bounded.

Notice however that the complement principle immediately predicts upward boundedness as well as apparent violations of upward boundedness (in case the related element of the matrix clause is itself moved like in the example below):

(26) *$[_{CP}$[Auf DEN Berg]$_i$ glaub ich nicht [daß Peter t_i gegangen ist]$_j$ [der Maria zu hoch war.]$_j$*
On the mountain believe I not that Peter gone is, which Mary too high was

Here the identifyee is moved out of an embedded clause to SpecCP of the matrix clause. One of the main empirical difference between the subjacency approach and the complement principle concerns sentences like (26): as soon as the identifyee is moved to a higher position the two approaches make different predictions. The complement principle predicts them to be good, the strict upward boundedness approach (derived via some version of subjacency or the ECP ruling out the intermediate traces) predicts them to be bad[14]. Again we find that the S-structure position of the identifyee determines the adjunction site. In (26) the identifier can be adjoined to the matrix

[14] The only possibility for the latter approach to predict those data to be good is to assume that extraposition takes place from the position the element the extraposed phrase is related to is moved to (cf. Ross 1967 for such a view). This would imply a strictly derivational and rule-ordering approach or that extraposition is last-cyclic, resp. a PF-phenomenon - an alternative we have already rejected (cf. section VI.1.4.2.).

clause (i.e. to CP) apparently violating upward boundedness, since the base position of the identifyee is in an embedded clause. The data under consideration favor the complement-principle approach as has already been mentioned by May & Guerón (1984) on basis of the following example:

(27) *$[_S$Which spy t does e_j [does Angelton believe $[_{S'}$ e_j that $[_S$ Burgess recruited e_j]] who ultimately became a mole$_j$]]* (G&M 1984, (32) p.16)

I conclude that some version of the complement principle is clearly preferable over any version of the subjacency account since it is both theoretically and empirically superior.

Before I will go into a more elaborate discussion on the various versions of the complement principle let me briefly discuss some analyses which seem to me to be the precestors of this sort of explanation.

2.1.1. Koster 1978

I have already given an overview of Koster's (1978) analysis and his arguments for base-generation (still with the assumption that we are dealing with an empty category - a trace to which the extraposed element is linked via some mechanism of Focus Linking). Now let me briefly present his account for the locality constraint. The problem he has to face in claiming that extraposition is constraint by an interpretive construal rule rather than a condition on movement is the following one: the alleged movement rules seem to be more island-sensitive than interpretive rules such as bound anaphora. This difference is exemplified by examples like:

(28) (a) *John said that a picture of himself was hanging in the office.*
  (b) **Who did he say that a picture of t was hanging in the office.*

Another problem is that in many languages (most) phrases of type XP are islands for movement rules, but not for rules of construal like Bound Anaphora (i.e. few languages allow for P-stranding but reflexivization into PPs is a well attested phenomenon). Koster accounts for this fact in claiming that the more limited distribution of empty nodes is not due to the fact that they are created by a different class of rules (movements), but simply to a property of the empty nodes. Empty nodes introduce essential complexities that are naturally absent in the case of lexically filled nodes (this approach anticipates more or less the core insight underlying the ECP). Koster posits the following rule:

(29) The Main Projection Rule
In core grammar, empty nodes are only possible in positions immediately dominated by categories of the main projection (i.e. V^i or S^j; $3 \geq i, j \geq 0$)

The MPR determines that all phrase nodes that are not directly dominated by the main projection line are too deeply embedded to be empty in core grammar. This sort of locality constraint is more or less parallel to the LCI assumed here, namely that the identifyee can be linked to an element which is directly adjoined to it or one maximal projection further up - anything else being too deeply embedded - i.e. too far away for Identification to be licensed. So empirically Koster's approach for the locality constraint and the one suggested here are similar. The difference being that Koster (1978) assumes (base-generated) empty categories that induce this locality constraint whereas I do not assume any empty categories left behind by extraposition. Rather the locality constraint is induced by the relation of Identification.

Moreover, as we have seen there is one crucial exception for this generalization - i.e. if the identifyee carries (contrastive) stress the locality condition becomes weakened and the identifyee can be more deeply embedded. Interestingly Koster discusses the possibility to extend certain locality domains - in this context he talks about core and marked phenomena. He observes the following cross-linguistic generalizations:

a) many languages have maximation strategies (e.g. pied piping)
b) few languages have P-stranding
c) underlined nodes may all be empty, but only a limited subset of the non-underlined nodes may be empty and if they are it is always at a cost.

It appears that phrase internal empty nodes occur only in maximally simple, accessible positions. Well-known complexity factors, like depth of embedding and nesting, govern the heavy restrictions on the distribution of empty nodes.[15] (He then tries to assimilate binding theory and ECP - a line of research which is still going on in the recent literature).

[15] Koster argues for the following complexity features:
Features of marked constructions:
- differentiation across languages
- variation in judgement
- susceptibility to lexical and non-grammatical complexity features

Marked structures, which are themselves complex in the sense of the evaluation metric are less stable and more sensitive to the disturbing influence of non-grammatical cognitive factors. Notice that this is exactly what we find in many of the data involving extraposition.

2.1.2. Asakawa (1979)

A similar locality constraint still stated in a rule-like (movement) account is given in Asakawa (1979):

(30) The element which is extracted out of NP is adjoined to the node which immediately dominates that NP.

Notice that this approach is rather similar to what I am proposing here, namely that the extraposed element is always higher than the element it is related to. This is a property of my analysis which differentiates it from any previous version of a *complement-principle account*. In the latter the extraposed element can or must be lower than the element it is related to. Let me just mention one of Asakawa's (1979) empirical evidence for his claim since it is also an argument for LCI and against the various other proposals suggested in the literature.

Asakawa deals with relative clause extraposition, extraposition from *'the claim'* as well as extraposition of PPs (exact for the last ones (i.e. PPs) all of them are instances of Identification). The most crucial piece of evidence concerns extraposition from subject position where it has been argued that the extraposed element attaches to VP (i.e. it is lower than the related constituent in the matrix):

(31) *The bricklayer appeared on TV who earned more money than a college professor*
(32) (a) *The gardener appeared on TV who earned less money than a carpenter, and the bricklayer did who earned more money than a college professor*
(b) **The gardener appeared on TV who earned more money than a college professor, and the bricklayer did, too.*

The following example involves an extraposed constituent which is related to the object of the matrix clause showing that it is in that case adjoined to VP:[16]

[16] I will discuss some more instances of VP-ellipses later in this chapter. There it will be shown that VP-ellipses cannot be taken as a straightforward test for the adjunction site of extraposed elements.

(33) (a) *I saw the bricklayer yesterday who earned more mone than a college professor, and my wife did, too.*

(b) **I saw the bricklayer yesterday who earned more money than a college professor, and my wife did, who earned less money than a carpenter.*

(34) *Sam picked those packages up which are to be mailed tomorrow last nicht, but he didn't want to do so until it had stopped raining.* (Ross, 1967)

(35) *I couldn't believe that there was a person under this roof who had already killed five times and who was preparing to kill again, although the detective told me with confidence that there was.*

(36) *I explained the fact to my son that money doesn't grow on trees, and my wife did, too.*

(37) *?John sent a photograph to the professor of the canal on the surface of Mars, and i think that Fred did, too.*

What is mainly important about Asakawa's analysis is that he already realizes the problem of positing a certain node for extraposition like for example Baltin (1978) has argued. Asakawa explicitly criticizes Baltin for the following rule:

(38) Move S' to the right bracket of S or VP (Baltin, 1978)

Explicitly mentioning S or VP is theoretically undesirable. However with Asakawa's locality constraint the landing site for extraposed constituents is automatically determined by the position of the correlated element. Let us now turn to the discussion of the various versions of the complement principle.

2.2. The Complement Principle

There are three main proponents of the Complement Principle: Guerón (1980), Guerón & May (1984) and C&R (1990) all of them differing in various important aspects. What they all have in common is that they capture the locality constraints on extraposed constituents by means of an interpretive relation between the element in the matrix and the extraposed constituent rather than via movement. Hereby they also capture some interpretive consequences of extraposition.

2.2.1. Guerón 1980

For convenience let me briefly recapitulate Guerón's main claim to get the locality constraint for extraposed constituents. Her analysis basically consists of two parts.

First, the complement principle is supposed to regulate the possible adjunction site for the extraposed constituent. Secondly, a Linking Rule which relies crucially on Topic-Focus relations makes sure that the extraposed constituent "finds" its appropriate antecedent.

(39) The complement principle:
The complement of X is a constituent governed by X (Guerón, 1980: 642)

Government is defined in the sense of Chomsky's Pisa lectures:

(40) Government
α governs β iff α minimally c-commands β (that is, α c-commands β and there is no γ such that α c-commands γ and γ c-commands β and not γ c-commands α)

(41) Focus Marking
(a) Mark the last argument in the c-command domain of the verb "Focus of S".
(b) Mark the VP "Focus of S"

(42) Complement Linkage
Mark the PP (or relative clause) to the right of S "Complement of Focus NP".

There are several problems with this account. First let me address the theoretical problems. What this analysis amounts to saying is that the antecedent has to be found (by a syntactic principle) somehow backwards. In a second step, if the appropriate antecedent is found the complement relation is established. It is not before this point that the relevant site for adjoining the extraposed element can be determined. Within my account the identifyee inherently determines the relevant adjunction site. There is no need for an additional rule of Focus Marking and Focus Linking.

The next problem for all the complement principle approaches is that it is at least not clear to me what a complement actually is. The complement principle is just a necessary but not at all a sufficient condition for being a complement (and it is far from being an exact definition for being a complement). If it would be a necessary and sufficient condition for an element Y to be a complement of another element X to be governed by X then every Y' would be a complement of SpecYP (notice that Guerón has to assume that maximal projections can be governors). It is not just any government relation that defines a complement relation but the complement relation has to be introduced as a primitive relation, not to be reduced to independently motivated principles or locality relations. Although it sounds pretty much established to speak of a complement relation it is not at all superior to our Identification relation,

which I argued to be independently necessary as a primitive relation in the theory of Grammar. Notice moreover that Identification is also a necessary component in licensing relations for empty categories. For example it has been argued in Rizzi (1990) that licensing of *pro* consists of two parts: a formal and an Identificational component. The same is true for various versions of the ECP. I have already argued that there are various ways for Identification to be instantiated. Here I was talking about Identification at domain D. Identification of empty categories might be a case of Identification at S-structure, whereas for example binding could be interpreted as an instance of Identification at LF (cf. section III.6.)

Comparing the assumption of Focus Linking with Identification there is an interesting point at issue. A prediction Guerón's analysis makes is that extraposition is only possible from focus NPs. On the other hand my analysis predicts a similar but still different set of data. I am claiming that extraposition is licensed via Identification and Identification is a relation that identifies an individual which has not been identified previously. Therefore the interpretation is somehow linked to being *new information*. New information is something which is clearly related to being the focus of a sentence, but it is not the same. Elements which are the focus of a sentence are a subset of elements which constitute "new information". Thus I do not predict that extraposition is just possible out of focus NPs. It turns out that the Identification approach makes the correct prediction. Consider under this light the following paradigm of question-answer pairs. (C&R take question-answer pairs to be the crucial test for focussed constituent). The following examples seem to confirm Guerón's prediction:

(43) Q: *Was habt ihr Maria verkauft?*
What have you Mary sold
A: *Wir haben Maria das BILD verkauft, das in der Auslage lag.*
We have Mary the PICTURE sold, which in the window lied

(44) Q: *Wem wurde das Bild, das in der Auslage lag, verkauft?*
Who was the picture, which in the window lied, sold
A: (a) **Wir haben das Bild MARIA verkauft, das in der Auslage lag.*
We have the picture MARY sold, which in the window lied
(b) **Das Bild haben wir MARIA verkauft, das in der Auslage lag.*
The picture have we MARY sold, which in the window lied

But if we change the relative ordering of the direct and the indirect object we see that her generalization is not borne out any more:

(45) Q: *Wem wurde das Bild, das in der Auslage lag, verkauft?*
Who was the picture, which in the window lied sold
A: *Wir haben MARIA das Bild verkauft, das in der Auslage lag.*
We have MARY the picture sold, which in the window lied

The example above tells us that we can have an extraposed relative that is not construed with the focus constituent of the matrix. The ungrammaticality of the answers in (44) indicates that the relevant generalization is along the following lines "a focussed constituent cannot intervene in the relation between the head NP and its relative". In fact this generalization seems to be correct as exemplified by the following data:

(46) Q: *Wer hat das Bild gekauft, das in der Auslage lag?*
Who has the picture bought, which in the window lied?
A: (a) *MARIA hat das Bild gekauft, das in der Auslage lag.*
Mary has the picture bought, which in the window lied
(b) **Das Bild hat MARIA gekauft, das in der Auslage lag.*
The picture has MARY bought, which in the window lied
(c) *Ich glaube, daß MARIA das Bild gekauft hat, das in der Auslage lag.*
I believe that MARY the picture bought has, which in the window lied
(d) **Ich glaube, daß das Bild MARIA gekauft hat, das in der Auslage lag.*
I believe that the picture MARY bought has, which in the window lied.

(47) Q: *Was hat die Frau gekauft, die eben im Geschäft war?*
What has the woman bought, who just in-the shop was
A: (a) **Die Frau hat das Bild von RENOIR gekauft, die eben im Geschäft war*
The woman has the picture of RENOIR bought, who just in-the shop was
(b) *Das Bild von RENOIR hat die Frau gekauft, die eben im Geschäft war.*
The picture of RENOIR has the woman bought, who just in the shop was

Another problem Guerón's analysis has to face is the fact that it is generally assumed that there is just one single focus in a given sentence. It would then be predicted that it is just possible to have only one extraposed constituent (if extraposition is just licensed by being linked to a focus NP). However, as we have seen in previous

chapters it is possible to have more than one element in extraposed position (obeying the nesting requirement, as a byproduct of the locality constraint). Guerón would have to admit two focussed constituents within one sentence. Even if this could be done it is still not clear anymore how complement linking should take place since then two elements would be marked as focus. Complement linking could not unambiguously determine the relevant complement relation. If one would stick to "one focus NP per sentence" then it is also not clear how the second extraposed constituent could be linked to its related element since then complement linking via focus could not apply anymore.

Moreover, Guerón cannot account for examples where the correlated element is more deeply embedded in another constituent. In all those cases the government relation between the two elements does not hold anymore. She would have to postulate some sort of LF movement of the "complement taker" in order to establish the complement relation. However, we have shown that it is S-structure rather than LF where the locality constraint holds. Notice that this is also a crucial point where my account differs from Guerón's since she claims that the complement principle holds at LF. As we will see any approach assuming LF to be the relevant level for the given locality constraint has to face the problem of why covert extraposition (i.e. further movement of the extraposed element at LF) is not possible, given that it should in principle be an instance of free application of move α. It should be possible to rightward move at LF. All the analyses (i.e. Guerón 1980 as well as Guerón & May 1984) suggesting that the complement principle is established at LF still have to assume that it is somehow the S-structure position of the extraposed element that is relevant.

On the other hand the account suggested here gets the interpretational properties by virtue of the assumption that Identification is a primitive relation at domain D, however in purely structural terms the output is checked at S-structure. This assumption however is also implicit in Guerón's analysis since she assumes that there is an additional level (where NP interpretation takes place) which has to do with discourse and topic-focus relations respectively. Guerón's analysis has crucial restrictions on 3 levels where we can dispense with one of those levels, just making use of S-structure and a level which has to do with the interpretation.

2.2.2. Guerón & May (1984)

The next analysis that makes use of a version the complement principle is found in Guerón & May (1984) although their account is slightly different. They particularly focus on the difference between relative (and complement) clause extraposition as opposed to result clause extraposition (however both of these relations are supposed to be complement relations - again the critique on the exact definition of complementhood arises). The main points of their analysis can be summarized as follows:

- Extraposition is derived via movement

Here it has to be noticed that especially in case of result clauses they do not give any real argument and it seems to me that again a base-generation approach would be preferable even within their approach for the following reason: the dependency relation (i.e. the complement relation and its locality restriction) is established (or rather checked) at LF - the level where interpretation is assumed to take place. At D-structure the relevant elements need not be in a head-complement relation (being a complement is not really a necessary lexical information). Therefore it is not necessary to postulate a trace, i.e. a basegeneration approach would be possible.

- The head-complement-relation has to be *reconstructed* at LF. This is done by means of quantifier raising (QR) of the head NP (to a position where it governs the complement in extraposed position).

- Extraposition from NP and extraposition from QP (i.e. result clause extraposition) differ in
  - a) the landing site of the extraposed constituent:
    - extraposition from NP attaches to S
    - extraposition from QP attaches to S'
  - b) the landing site of QR of the head NP/QP in the matrix clause
    - NP undergoes QR adjoining to S
    - QP undergoes QR adjoining to S'

G&M assume that the relevant landing site of extraposition derives from the complement principle (holding at LF) since the relevant head has to undergo QR and

the relevant heads differ in their landing site[17] - the output however has to result in a configuration where the head governs the complement[18]. A similar criticism as the one briefly mentioned in the previous section is at stake here. There is nothing to prevent LF extraposition in order to fulfill the complement principle. Thus at S-structure the relevant adjunction site has to be already determined. Therefore the difference in adjunction site at S-structure is a construction specific statement, which is meant in order to capture the empirical data. However it is not the case that the S-structure position of extraposition is derived by any independently motivated principle.

The second point of criticism has to do with the main spirit of this account, namely that the complement principle is fulfilled at LF by virtue of QR. This phenomenon is motivated by assuming that the complement relation has to be reconstructed at the level where interpretation takes place. However, under standard assumptions reconstruction is a different phenomenon. An element which is overtly moved can behave as if it would still sit in its base-position for some phenomenon (e.g. binding). Especially since G&M assume that extraposition is derived via movement, nothing in principle would prohibit reconstruction in this sense. The complement principle could easily be fulfilled by reconstructing the moved (extraposed) element in the first place (presumably followed by QR of NP pied piping it along). This then creates the same problem mentioned before: the complement principle (holding at LF) cannot really determine the landing site of the extraposed element (at S-structure). In addition to the problem that there is no principled reason to exclude further rightward movement of the extraposed element at LF there is also an empirical evidence showing that they even have to assume the possibility to 'extrapose' at LF. Consider the following sentence with its S-structure representation:

(48) *$[_{CP}$ $[_{VP}[_{DP}$viele Bücher gesehen$]$ $[$die er nicht lesen konnte$]]_{VP}$ $[_{C}$hat Peter gestern$]]$*

[17] Concerning this assumption they do not give any independent reason or motivation for this claim. The only reason they give has to do with the definition of variables, which is by no means an undisputable assumption (but I will not go into a detailed discussion here).

[18] The relevant definitions are given below (I will not discuss them in more detail):

(i) *In a sequence of categories $\alpha_i, \beta_i', \ldots \beta_i''$ in a structure σ, $\beta_i', \ldots \beta_i''$ are complements to α_i, only if α_i governs $\beta_i', \ldots \beta_i''$.*

Assuming that S', NP, VP, AP and PP are maximal projections, government is defined as below:

(ii) *α governs β $=_{df}$ α, β are dominated by all the same maximal projections, and there are no maximal projection boundaries between α and β.*

They furthermore assume that the choice of β in (i) is limited to maximal projections. The choice of α however may vary - establishing different kinds of complementation relations.

As the sentence in (48) shows, it is a possible option for VP-topicalization to move the 'extraposed' relative clause along as long as the coindexed element of the matrix is within the VP. This imposes several problems for G&R. First (48) clearly shows that it is the S-structure position of the head NP that is relevant for determining the adjunction site. Under their approach the relative clause in (48) is not in a position where the complement principle can be fulfilled at LF. In addition to that this sentence also suggests that they have to assume the possibility for LF-extraposition for the following reason: The (head)-NP *viele Bücher* occurs within the VP that occupies SpecCP. Therefore QR has to take place from a right branch. Moreover recall that G&R assume that both the head NP and the relative clause adjoin to S (i.e. IP). This is a crucial assumption for deriving the differences between extraposition from QP. This means that in (48) the NP has to lower to adjoin to IP. If they would allow for adjunction of NP to S' (i.e. CP) in this case then there is no obvious reason to have the restriction they postulate in the first place. If it would however adjoin to IP then the trace of the NP would not be c-commanded. In order to fulfill the complement principle the relative clause that is adjoined to VP (i.e. extraposed) would have to adjoin to IP at LF, too. This indicates that 'extraposition' at LF must be a possible option. If it is possible in (48) then there cannot be any principled reason to exclude LF-movement of the 'complement' to a position where it is governed by its head at LF in all other cases as well. A potential possibility to solve this problem would be to claim that before QR (and LF-extraposition) takes place the fronted VP is reconstructed to its base-position. If so then we would overcome the problem of LF lowering of the QRed NP. However, in order to fulfill the complement principle the relative clause would still have to move at LF to a position where it is governed by the QRed head.

Under both theoretical possible options LF-extraposition is a necessary assumption. Notice that the second choice (where the VP is reconstructed) is creating another problem for Guerón & May. If this would be a possible derivation then there is no way to explain the ungrammaticality of VP topicalization leaving the head of the complement relation behind at S-structure:

(49) (a) **gesehen, die er nicht lesen konnte, hat Peter gestern viele Bücher*
seen, who he not read could, has Peter yesterday many books

(b) *gesehen hat Peter gestern viele Bücher, die er nicht lesen konnte*
seen has Peter yesterday many books, who he not read could

Under G&M's analysis the S-structure position of the head NP should not make any difference for the grammaticality of the minimal pair above, since in all of the relevant sentences the complement principle could equally be fulfilled at LF.

Another problem is again the stipulative character of the obligatory rightness of extraposition. There is nothing at all in this account to prevent the element which is extraposed to left-adjoin to the position it can be right-adjoined to. Thus the analysis in G&M does not overcome the problem of construction specific statements concerning extraposition.[19]

2.2.3. Culicover & Rochemont (1990)

Although C&R's analysis differs in various respects of the previous analyses overcoming some of the problems discussed above it still has some theoretical and empirical problems. The main points where C&R depart from the analyses presented so far are summarized below:

- Extraposition does not involve movement. There is no trace left.
- The complement relation is primitive relation - not to be equated with theta-marking relations (as it is for example assumed in G&M).

C&R argue for the latter property partly on basis of Baltin's (1983) already discussed criticism that the Complement principle is required to hold only between lexically filled maximal projections, whereas theta marking may also hold between empty elements. However, they conclude that the complement principle is independently motivated by predication. They argue that the complement principle

- makes movement redundant (the principle of full interpretation is satisfied by virtue of the complement principle)
- explains the subjacency effects
- is the only way to license extraposed elements (which are neither arguments nor predicates nor operators)

All this latter properties of their analysis are similar to the proposal suggested here the only difference being that I am not using the complementation relation but Identification (which as I argued to be independently motivated). Moreover, the criticism already mentioned concerning the nature of the complementation relation also

[19] Notice that 'upward boundedness' receives an interesting interpretation in their account. QR is independently assumed to be clause bound. Extraposition cannot adjoin to a higher clause since QR is clause bounded. Therefore the extraposed constituent will never move beyond the clause its antecedent occurs in.

applies here. It is still not clear what a complement really is. C&R give the following description of a 'complement':

> *'the 'complement' to a maximal projection is a phrase that holds an adjunct or argument relation to the head of the phrase it takes as antecedent.*
> (C&R 1990, Fn4: 26)

Again being a complement is both structurally defined (by virtue of the complement principle) and w.r.t. the relation between the two elements. However, a precise definition of being a complement is still missing. It is still not clear how it is determined whether an element can take another element as its complement-taker.

An additional difference to the previous analyses is that it is assumed that the complement principle holds at S-structure (rather than LF). This is similar to what I am proposing and thus overcoming the problems discussed above w.r.t. G&R's analysis. One piece of empirical evidence they give is the following sentence[20]:

(50) (a) **Who told her$_i$ that Sam was taking a student to the dance t_{CP} that [the teacher]$_i$ liked]?*
(b) **Who told her$_i$ that Sam was taking [which student]$_j$ to the dance t_{CP} that [the teacher]$_i$ liked]$_j$?*

According to standard assumptions about wh-in situ phrases they are moved at LF to SpecCP. If the complement principle would not be fulfilled before LF, then the extraposed constituent would necessarily be adjoined at the matrix IP or higher at S-structure on par with the following example, where the wh-phrase has been moved overtly:

(51) *How many girls did t_{IP} he$_i$ invite to the party]$[_{CP}$ that John$_i$ dated in high school]*

[20] Let me just mention an other piece of evidence that could suggest that the complement principle holds at S-structure rather than LF. This argument is due to Baltin (1987). Apparent antecedent-contained deletion of VP in the examples below would be better analyzed as involving VP ellipsis within an extraposed complement adjoined to VP since the assumption that there may be no antecedent-contained deletion of VP can explain the lack of ambiguity in (ii):
(i) *Bill hit the man who asked him to*
(ii) *Who thought that Fred read how many of the books that Bill did?*
The sentence in (ii) can just have the reading given in (a) below but not the one in (b):
(a) *Who thought that Fred read how many of the books that Bill read?*
(b) *Who thought that Fred read how many of the books that Bill thought he read?*
However, C&R notice that this argument is not unproblematic. See also Diesing (1992) for a criticism.

Now let us turn to the last important difference between C&R's approach and the previous analyses. This difference concerns the definition of the complement principle itself and the locality condition for extraposed constituents. The formulation of the complement principle they propose is given below:

(52) β is a *potential complement* of α (α,β = X^{max}), only if α and β are in a government relation.[21]

This definition of the complement principle is both stronger and weaker than the ones that were previously proposed. As they point out, it is stronger since it requires a c-command relation between the two elements, which was not the case in the previous formulations. On the other hand it is weaker since it does not require that the 'antecedent' governs the extraposed element - it is enough for the complement relation that there is a government relation, no matter which element governs which.

Let me briefly mention the theoretical problems with such a view. Since I am assuming that relations have to be asymmetric this weaker formulation seems to me problematic since it implies a potential symmetry. Notice that my analysis involves an apparent symmetry, both elements seem to depend on each other - but I concluded that this apparent symmetry is indeed to be analyzed as involving two asymmetric relations (as discussed in chapter I.) A second problem is the stipulative character of the obligatory rightness of extraposition. Again there is no way to make it follow from any principle that the extraposed constituent has to right-adjoin rather than left-adjoin. We still find a construction specific statement.

Now I will turn to some empirical problems. Moreover I will compare the analysis I am proposing with C&R's formulation of the complement principle. It is a result of their proposal that the extraposed element can either be hierarchically higher or lower than its antecedent whereas in my account it has to be higher. Notice that under my account the nesting effect is immediately predicted whereas they have to introduce an additional (interpretive) constraint. My analysis captures more data with one principle.

[21] For completeness let me give the definition of government they are assuming:

(i) α *governs* β if α c-commands β and there is no Δ, Δ a barrier for β, that excludes α.

(ii) Δ is a *barrier* for β iff (i) Δ is an X^{max} that dominates β and (ii) Δ is not theta-governed (directly theta-marked).

(iii) Δ (Δ = X^{max}) *excludes* α if no segment of Δ dominates α.

(iv) Δ (Δ = X^{max}) *dominates* α only if every segment of Δ contains α.

The reason for allowing both options is mainly based on data involving VP-ellipsis. Their claim is that the extraposed element is adjoined to VP if it is related to the object and adjoined to either VP or IP if it is related to the subject. According to C&R, in this case it can be lower than its antecedent. Since the analysis I am proposing here does not allow for an identifier to be lower than the identifyee I have to address the apparent evidence for VP-adjunction of a subject-related identifier. I will devote the next section to a discussion of VP-ellipsis in order to justify the analysis given here and also to show that VP-ellipsis cannot be used as a reliable test for the adjunction site of extraposition.

2.3. VP-ellipsis and the Locality Constraint on Identification.

There are two different paradigms discussed in C&R which are intended to show that subject-related relative clauses can be adjoined to VP in which case they would be lower than their antecedents. First consider the following paradigm apparently showing that extraposed elements which are construed with subjects can be adjoined to VP (rather then IP):

(53) *A MAN came in with blond hair, and a WOMAN did [e] TOO.*

(54) *Although none of the MEN did several of the WOMEN went to the concert who were visiting from Boston.*

(55) (a) *A MAN came in with BLOND hair, and a WOMAN did [e] with BROWN hair.*

(b) *Although none of the MEN did who were visiting from NEW YORK, several of the WOMEN went to the concert who were visiting from BOSTON*

(C&R 1990: 30)

These examples are intended to show that when extraposition is from subject-position, VP ellipsis apparently may but need not take the extraposed constituent. This behavior is argued to be expected under their assumption: the extraposed constituent is analyzed as being adjoined to VP. Then VP-ellipses is assumed to apply to either node of the two VPs created via adjunction. Notice that C&R mention the crucial fact that for these sentences to be well-formed they have to involve some contrastive stress on the relevant terms. Without giving an explanation for this phenomenon they claim that it should not make a difference w.r.t. the adjunction site, i.e. they claim that this fact does not undermine their point that extraposition from subject position can adjoin to VP. This fact is however crucial and it follows from the analysis I will propose for

these sentences.

The second piece of evidence (again involving VP-ellipsis) is as follows. If it can be shown that a particular constituent C must be attached to VP and if a subject related relative clause in extraposed position can precede this constituent then it follows that the extraposed subject relative clause must be attached to VP as well (i.e. it must at least be lower than the constituent that follows it). The following data show that an element that is arguably adjoined to VP follows an extraposed subject relative clause. In (56) we find an adverbial clause that is analyzed as being adjoined to VP. C&R again use VP-ellipses as well as VP-topicalization in order to show that those adverbials are indeed adjoined to VP:

(56) (a) *Mary came into the room as quickly as possible, and John did too.*
(b) *They asked him to come into the room as quickly as possible and come into the room as quickly as possible, he did.*[22]

On basis of the above mentioned argumentation, i.e. that a (right-adjoined) element preceding another (right-adjoined) constituent must be lower (i.e. as well being adjoined to VP) the following examples are presented in C&R:

(57) (a) *Some women came in (who were) from Chicago as quickly as possible.*[23]
(b) *A man came into the room that Mary recognized as quickly as he could.*

In addition it is shown that VP-ellipsis may include both the adverb phrase and the extraposed constituent:

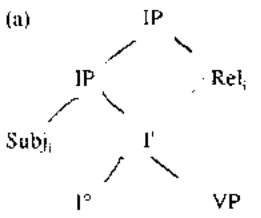
[22] The second piece of evidence for assuming that the adverbial is adjoined to VP involves condition C effects:
(i) **He, came into the room as quickly as John, could*
It is not possible to construe a subject pronoun in the matrix as being coreferent with an R-expression in the adverbial indicating that the pronoun c-commands the R-expression. I have already addressed the problematic aspects of condition C-effects as a test for the adjunction site in section III.4.3. I have argued that it cannot be taken as a reliable test.

[23] It was pointed out to me by Edwin Williams (p.c.) that the sentence in ((57a) does not sound perfectly well-formed. This is in accordance with my analysis. Moreover it suggests again that Identification by a PP (i.e. a non-coindexed constituent) is a different phenomenon, that I have not addressed.

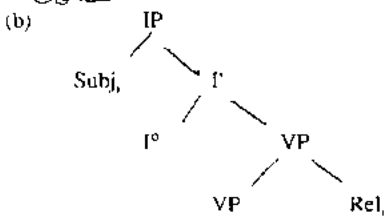
(58) (a) *Some women came in (who were) from Chicago as quickly as possible and some men did too.*
(b) *A man came into the room that Mary recognized as quickly as possible, and a woman did too.*

Let me summarize the problem we have to address on basis of the examples given in C&R. According to my analysis any extraposed constituent related to a constituent in the matrix clause has to be higher as its correlate (i.e. it has to be adjoined to the first maximal projection dominating the identifyee (59a)). LCI predicts that a subject-relative clause is adjoined to IP if the subject occurs in IP whereas C&R argue that it can also be adjoined to VP as indicated in (59b) below:

(59) (a)



(b)



If VP-ellipsis is analyzed as deletion of VP the structure in (59a) obviously creates a problem. If the extraposed element is deleted along with the rest of the deleted constituent (VP) then obviously it cannot be adjoined higher than the subject NP before deletion has taken place. My analysis can only account for the cases where the subject related constituent remains. Notice that this seems to be the unmarked case since then no special stress is needed. LCI cannot account for the examples where the subject relative clause is apparently deleted along with the VP, the subject NP being still overtly present. The subject itself would have to be deleted as well since it is lower and thus dominated by the constituent to be deleted.

The problem we are confronted with here reflects the general problem to determine the relation between examples like:

(60) A: *Who can do it?*
B: *John can.*

Any analysis for this phenomenon has to capture the fact that the VP in B is the same than the VP in A. In general there are several possible ways to analyze VP-ellipsis and

its properties. There is the 'deletion'- analysis discussed in Ross (1967) where at some level of representation B includes a full VP (*John can do it*). The rule of VP-deletion applies to this representation. Deletion can take place under identity with some other VP (the general restriction being *recoverability of deletion*). Interpretation here would take place prior to deletion. In a principle and parameters framework assuming the T-model and LF as the relevant level of interpretation this would translate into claiming that VP-deletion is a PF-phenomenon. Thus at LF the full VP would still be available since PF-processes do not have any reflex at LF.

The alternative (interpretive) proposal (e.g. suggested in Williams (1977)) is to say that B has an S-structure representation where the VP is occupied by a phonologically null element which is anaphorically related to the VP in A. The empty element receives an interpretation by means of the anaphoric relation.

2.3.1. A possible solution

A strictly interpretive account for the facts under discussion has to capture the fact that the empty VP can be anaphorically related to an apparently discontinuous constituent (i.e. the VP and the extraposed constituent). In other words we have to give an explanation for when and why the property the relative clause denotes can be understood as a property of the subject of the second conjunct even if the relative clause is not overtly present (i.e. apparently deleted). Under such an approach it would not be a matter as to whether the relative clause is adjoined to the phrasal constituent that is deleted (at least before PF) but a matter as to whether an NP can be interpreted as having the same property as an NP in the first conjunct. This amounts to saying that the property in the relative adjoined to the first conjunct can be understood as identifying two independent NPs. This possibility is independently attested by the existence of *split antecedents* (discussed in chapter VI). We have seen that split antecedents are apparently only possible in coordinated structures. Notice that the examples above crucially involve coordinated structures. Thus the existence of split antecedents can be used as a first evidence for a strictly interpretive account.
However, the present data, involving stress on N still behave somehow differently, i.e. there is something else going on here. It seems that in case of stress on N the possibility to identify two independent constituents is not contingent on VP-ellipses. Consider the following sentences:

(61) (a) *Ein MANN, der blonde Haare hatte, traf gestern eine FRAU.*
a MAN, who$_{masc}$ blond hair had, met yesterday a WOMAN.
(b) *Ein MANN traf gestern eine FRAU, die blonde Haare hatte.*[34]
a MAN met yesterday a WOMAN, who$_{fem}$ blond hair had.

Both in (61a-b) there is an interpretation available where both individuals (i.e. the woman and the man) are interpreted as having blond hair. This example indicates that stress on N makes it possible to have two interpret two separate NPs 'identified' by the relative clause. Both individuals denoted by the two NPs under consideration are interpreted as having the property denoted in the relative clause. We will see that the analysis for contrastive stress presented in IV.7.2. can account for this phenomenon.

A second piece of evidence for the claim that it is necessary to assume that the interpretation of VP-ellipsis can involve *discontinuous constituents* is the following paradigm. Here this phenomenon is independent of extraposition. The data are discussed in Prüst, Scha & van den Berg (1994):

(62) (a) *Maaike likes belly-dancing*
(b) *She hates waltzing*
(c) *Saskia does too.*

The sentence in (62c) is ambiguous. Under one reading the elliptical VP takes the VP in (62b) as its antecedent (where *Saskia hates waltzing*). The other interpretation results from anaphorically relating the VP in (62c) to both (62a) and (62b) (where *Saskia hates waltzing and likes belly dancing*). This phenomenon indicates that it cannot simply be deletion of a particular constituent that is responsible for the interpretive properties attested in (62). Here we have two (heavily) discontinuous constituents that can be used as the antecedent for the elliptical VP: they can even occur in two separate sentences. On basis of these examples Prüst, Scha & van den Berg (1994) conclude that the problem of VP-anaphora must be viewed from a general discourse perspective.

[34] Notice that the ungrammaticality of the following sentence does not cause a problem for the present analysis:
(i) *Ein MANN traf gestern eine FRAU, der blonde Haare hatte.*
A MAN met yesterday a WOMAN, who$_{masc}$ blond hair had
The relative clause cannot 'directly' identify the NP *ein Mann*, i.e. the relative pronoun has to agree in gender with the second NP *eine FRAU*. This is expected since we have already seen in section 2.2.1. that a focussed constituent cannot linearly intervene between the identifyee and the identifier. In (i) the second NP *eine FRAU* would linearly intervene.

Moreover, consider the following paradigm that involves the same problem as the data we have to give an account for.

(63) (a) *John jogs in the street in the morning*
(b) *(and) Mary does in the park*
(64) (a) *John jogs in the STREET in the morning*
(b) *(and) Mary does in the PARK*
(65) (a) *John jogs in the street in the morning*
(b) *and Mary does in the evening.*

First it is interesting to notice that according to Prüst, Scha & van den Berg (1994) some speakers prefer (64) over (65). In (64) we find contrastive stress. The preferred reading in (64) is that *Mary jogs in the park in the morning* (besides another less preferred interpretation where *Mary jogs in the street in the park*). These cases are parallel to our examples but without involving extraposition. In this case it seems to be unavoidable to allow a discontinuous constituent to count as the antecedent for anaphoric relations. Descriptively in this case the VP-anaphor under consideration occupies one or more deletion sites:

(66) (a) *John jogs in the street*
(b) *and Bill does ∅ too.*
(67) (a) *John jogs in the street in the morning*
(b) *and Mary does ∅ in the park ∅.*

On basis of these examples (among others) Prüst, Scha & van den Berg (1994) come to the conclusion that:

> *"an adequate treatment of VPA requires the whole anaphoric clause to be matched with its antecedent. Complex embedded VP anaphora in ambiguous quantifier structures induces clausal parallelism with regard to quantifier and binding structures. Furthermore it is necessary to establish parallel and non-parallel elements, both of which may be located outside the VP. [...]*
> *the semantic relation between elements of the anaphoric clause and the elements of the antecedent strongly influences the interpretation of the anaphoric clauses in these examples. [...] As in the case of utterances that exhibit Gapping, these instances of VPA cannot be explained by some prefabricated structure with holes/variables in it: neither the number, nor the types nor the places of the gaps (the deletion sites) can be specified beforehand, they have to fall out of the interpretation process..."*
>
> (Prüst, Scha & van den Berg 1994: 323)

I take their evidence to be convincing enough to establish an analysis along their lines, i.e. allowing for discontinuous constituents to function as antecedent for VP-anaphora. Moreover, the paradigms discussed above suggest that the possibility for apparently identifying two independent NPs with only one identifier is neither contingent on extraposition nor is it contingent on VP-ellipsis. As the data discussed so far indicate, the crucial factor seems to be stress on N.

A final argument against a simple deletion or copying theory is the fact that the relevant constituent in the second conjunct can contain different features as the one in the first conjunct (thus we find a violation of the strict identity requirement). This involves for example different phi-features as in:

(68) *Eine FRAU kam herein, die blonde Haare hatte und ein MANN auch.*
A woman came in, who blond hair had, and a man too.

German relative clauses contain relative pronouns which have to match in phi-features (i.e. gender and number) with their head NPs. If the two antecedents differ in gender the relative pronoun would have to be different. It cannot be argued that it is really the very same constituent which is copied (or deleted) that is responsible for the correct interpretation. If it would be we would get:

(69) *Eine FRAU kam herein, die blonde Haare hatte und ein MANN kam herein die blonde Haare hatte.*
A woman came in, who$_{fem}$ blond hair had, and a man came in who$_{fem}$ blond hair had.

A parallel phenomenon is discussed in Williams (1977) concerning different Tense features:

(70) A: *Bob left.*
B: *Bill will, too*

These data show that certain features can differ without violating the identity requirement.

Moreover, it is interesting to mention that Williams (1977) assumes a rule of Discourse Grammar to be responsible for VP-ellipsis in general. This has to do with the fact that VP-ellipsis - as already shown - can take discourse antecedents, i.e. the relevant antecedent need not be in the same sentence (and moreover the anaphoric

relation can violate several locality constraints which hold for other relations, like e.g. the complex NP constraint or the coordinate structure constraint). However, Williams (still assuming a derivational system) assumes a level of discourse representation to be ordered after the level of LF where sentence interpretation takes place. I argued that it is domain D where such phenomena should be captured.

I will now give a brief sketch (although much less elaborated than the account given in Prüst, Scha & van den Berg) of a possible way to look at the apparently problematic data involving extraposition. On basis of the data discussed so far I concluded that stress on N is the crucial property allowing for one identifier identifying two different NPs. Consider again the analysis we gave for the interpretation of a relative (or 'complement' clause) identifying an NP which bears contrastive stress on N like in the example below:

(71) *Peter hat das AUTO gekauft, das bei Egon in der Werkstatt lag, und nicht das Motorrad.*
Peter has the CAR bough, which at Egon in the garage stood, and not the motor bike.

I have argued that in this case we are talking about a particular set of individuals that is further restricted. In this case the superset is identified by the relative clause rather than the actual discourse referent. I have argued that in case of contrastive stress on N the identifier is necessarily associated with X (i.e. the index of N). Furthermore I said that this identified X corresponds to the set of alternatives to which the denoted DP is contrasted. Moreover we have seen that the presupposition we arrive at is as follows: there are other individuals satisfying the property denoted by the identifier. In other words Identification of X by the relative clause provides a set of individuals that have the same property (namely the one denoted by the relative clause). Contrastive stress indicates that the property denoted by N is contrasted to some other possible properties the individuals of the established set have in common. For (71) this means that DP denotes an individual whose distinguishing property is *being a car*. The set of alternatives has one particular property in common, that is all of them are assumed to have the property of *standing in Egon's garage*. Therefore the second DP occurring in the example above (i.e. *ein Motorrad*) is also interpreted as having this property, which is denoted by the relative clause. (Notice that (71) is also a case where one relative clause can be associated with two different NPs. In this case this phenomenon is again neither contingent on extraposition nor on VP-ellipses. The only relevant property seems to be contrastive stress on N.)

Given the analysis of contrastive stress in interaction with identification we can derive C&R's facts without having to say that the relative clause is adjoined to VP. Consider again the relevant sentence repeated below for convenience:

(72) *Eine FRAU kam herein, die blonde Haare hatte und ein MANN auch.*
A WOMAN came in, who blond hair had, and a man too.

The relative clause adjoined to the first conjunct identifies the index X, which is associated with the DP *eine FRAU*. This means that the superset (i.e. the set of alternatives) is interpreted as having the property denoted by the relative clause in common. The second DP *ein Mann* involves contrastive stress on N as well. This means that it has to be contrasted to a set of alternatives as well. There is a set of alternatives available, i.e. a salient set of alternatives. This is the very set of alternatives that the previously mentioned individual was contrasted to. This superset is restricted by the relative clause. Therefore the interpretation arises that all the members of this set have the property denoted in the relative clause in common. The second individual under consideration is most naturally assumed to being picked out of the same set of alternatives. Therefore the interpretation arises that *ein MANN* is interpreted as having the property denoted in the relative clause. It is not necessary to assume that the relative clause is adjoined to the second conjunct before deletion takes place. The set of alternatives is simply salient enough by virtue of being associated with a DP in the first conjunct.

This means that we can explain that this phenomenon is not depend on VP-deletion. Rather, it is depend on the discourse salience of a property of individuals within a given set (the set of alternatives). We predict that the same effect can also occur inter-sententially. This prediction is confirmed by the following discourse:

(73) *Es gibt einfach zu viele Menschen mit schwarzen Lederjacken. Gestern kam da ein MANN ins Lokal der eine schwarze Lederjacke trug. Tja und später kam dann sogar noch ein kleines MÄDCHEN.*
'There are simply too many people with black leather jackets. Yesterday a MAN came into the pub, who was wearing a black leather jacket. Well, and later even a little GIRL came in.'

One possible interpretation for the last sentence within the example in (73) is that the little girl was also wearing a leather jacket. Here it cannot be argued that deletion is responsible for this interpretation. Rather it is more an implication which can be drawn from the discourse salience of *individuals wearing black leather jackets*. The new

discourse referent (*a little girl*) can be interpreted as being a member of this set of individuals. This explanation can also explain the following effect. Consider again the example in (55) where we find contrastive stress on N but a second relative clause is still present. This example is repeated below for convenience

(74) (a) *A MAN came in with BLOND hair, and a WOMAN did [e] with BROWN hair.*

(b) *Although none of the MEN did who were visiting from NEW YORK, several of the WOMEN went to the concert who were visiting from BOSTON*

If a second relative clause (the identifier) is overtly present it is necessarily the case that some element within this identifier is (contrastively) stressed. This means that the (salient) common property of individuals must be contrasted to some other property of another set of individuals. We can also explain the following example again taken from C&R, where a rather stipulative account is given to explain this fact:

(75) *A MAN with blond hair came into the room, and a WOMAN did TOO.*

(75) allows for a reading where the NP of the second conjunct is interpreted as having the same property as the subject of the first conjunct. In order to justify the claim that VP-ellipses (deleting the extraposed VP-adjoined PP along with the VP) shows that extraposition from subject can adjoin to VP C&R have to assume that (75) is derived from the sentence in (76). They assume that the PP within the second conjunct must occur in extraposed position before VP-ellipsis takes place:

(76) *A MAN with blond hair came into the room and a WOMAN came into the room with blond hair TOO.*

The assumption of 'vacuous' extraposition (which is of course a necessary assumption for C&R) creates several problems. Consider the following sentence. Here again two apparent different sites of 'deletion' occur:

(77) *Peter ist gestern nachmittag SCHNELL ins Zimmer getanzt und Maria LANGSAM*
Peter is yesterday afternoon quickly into the room danced and Mary slowly

Here C&Rs line of reasoning cannot be maintained since the adverbial element *schnell* can not occur in extraposed position. In general these adverbials seem too 'light' to be extraposed in German. This is exemplified by the following ill-formed sentence:

(78) **Peter ist gestern nachmittag ins Zimmer getanzt SCHNELL*
Peter is yesterday afternoon into the room danced quickly

C&R have to allow extraposition (before deletion) also for elements that are not possible in extraposed position otherwise. Notice that they have to admit (in a footnote) that

> *"in order for (23)[=76] to be derived from 25[=77], the principles licensing the interpretation of an ellipted VP must make reference to the interpretation of the entire sentence containing the antecedent VP."*
>
> (C&R 1990 Fn.22: 32)

They have to assume some interpretational device in addition to VP-deletion. Therefore I take the account suggested here as being superior to C&R's analysis. First, it captures the data with simply one mechanism (whereas C&R need two different devices). Secondly we also have an explanation as to why contrastive stress is a necessary prerequisite for the interpretation we arrive at (C&R however cannot explain this fact).

There is one problem left for the present analysis. We have said that the possibility for identifying two discontinuous NPs with only one identifier is neither contingent on VP-ellipses nor on extraposition. (Notice that we have already seen empirical evidence supporting this view). The following sentence seems to contradict this prediction.

(79) *A MAN with blond hair came into the room, and a WOMAN came into the room TOO.*

The sentence in (79) does not allow for the reading that the woman had blond hair, too. This could suggest that 'adjunct deletion' is really a function of VP-ellipses. However, I think that this might be true in the case above but for different grounds than C&R claim. Notice that the sentence above differs from the ones we have been used as evidence that there is no such contingency of VP-ellipsis and 'adjunct-deletion'. In (79) we find a coordinated IP. The reason for why it is not possible to interpret the NP (a woman) above as *'a woman with blond hair'* could be as follows. If the whole VP is repeated, then the relevant information about the woman is *'that she also came into the room'*. In the case above the PP *'with blond hair'* could be argued to not restricting the set of alternatives. It is rather interpreted as an additional information

(more or less a nonrestrictive modification for the man[25]) the relevant information is that a particular individual (who had blond hair) came into the room (as opposed to some other individuals who didn't come in). Notice that Prüst, Scha & van den Berg come to a similar conclusion. They assume that

> *'mostly only non parallel elements in the antecedent contribute to a full interpretation of the anaphoric clause'.* [26] (Prüst, Scha & van den Berg, 1994)

What this amounts to saying is that if the second conjunct does not have a parallel then the element under consideration is interpreted along with the ellipted element no matter whether this is a full constituent or a discontinuous one. However if the relevant element is repeated then the interpretation of a constituent in the first conjunct is not available anymore. This can also be seen on basis of the following example:

(80) (a) *John jogs in the street*
(b) *Bill does too*

In (80) *'in the street'* does not have a parallel in the second conjunct. It contributes to the full interpretation of (80b). I take it that the explanation for this fact can also give an explanation for the generalization in C&R that the relevant VP-ellipsis facts are contingent on extraposition.[27]

Whatever the exact conditions for the interpretation of conjunction, VP-ellipses, contrastive stress etc. are (which goes beyond the scope of the present discussion) I conclude that the data used by C&R cannot be taken as evidence that extraposition from subject-NPs can adjoin to VP. Moreover I have argued that there is another

[25] The same is true for the following sentence:
(i) *A blond-haired MAN came into the room, and a WOMAN did TOO.*
Again the adjective has to be interpreted as not restricting the set of alternatives, but as nonrestrictive modification for the individual *'a man'*.

[26] 'Parallel' here is interpreted as being 'parallel in thematic roles'. Cf. Williams (1994) for a similar conclusion.

[27] Notice that this line of reasoning does not imply that the following sentence is well-formed:
(i) *John saw a man who laughed and Mary did who cried*
In that case there is not possible adjunction site available for the relative clause in the second conjunct. There has to be an overt identifyee present.

account available (which I could just sketch) which captures even more data.[28] There is no reason left to assume that the identifier can adjoin to a position lower than the identifyee.

3. Müller (1994)

Since it is one of the most recent approaches to deriving the peculiar difference between left- and rightward relations I will discuss Müller (1994) in more detail. What can be said in advance is that he does again face the problem that he has to assume a genuine difference between left- and rightward movement - a view I have already argued to be problematic.

Let me briefly introduce the main claims of his analysis. Then I will discuss the theoretical and empirical problems for such an approach. Müller assumes that the asymmetry between left- and rightward movement derives from the prohibition against improper movement. This is formally implemented as the so called *Principle of Unambiguous Binding* (cf. Müller & Sternefeld (1993)):

(81) Principle of Unambiguous Binding (PUB):
A variable that is α-bound must be β-free in the domain of the head of its chain. (where α and β refer to different types of positions).

This principle states that movement to an A-bar position of a certain type α may never be followed by movement to another type of position β. Different types of positions are of course A and A'-positions, as well as different A'-positions derived via wh-movement, topicalization or scrambling respectively. With this principle Müller & Sternefeld (1993) try to give an attempt for some instances of improper movement, strong cross over etc. It was originally proposed to derive the ill-formedness of the following representations:

(82) **[A man]$_i$ seems [$_{CP}$ t' (that) [$_{IP}$ there was killed t$_i$]]*

Successive cyclic superraising is excluded since the (contextually determined) variable t$_i$ is (ambiguosly) bound by the intermediate trace t', in the embedded SpecCP position

[28] Notice that the second evidence C&R give as evidence that subject-related relative clauses can adjoin to VP has to do with VP-ellipses and the interpretation of VP-adverbs. Without discussing this further I assume that a similar account as the one just sketched can be given for these data.

AND by the chain antecedent *a man*, in the matrix SpecIP position.

PUB also rules out successive cyclic long distance scrambling. In this case the scrambled element is first moved to an A'-position of a certain type (SpecCP) and then this movement is followed by A'-movement of another type (adjunction to VP, i.e. scrambling):

(83) **daß niemand $[_{VP}$ Pudding$_i$ $[_{VP}$ sagt $[_{CP}$ t'_i daß sie t_i mag$]]]$*
that nobody jelly says that she likes

Given this principled account for ruling out improper (A'-movement) Müller then tries to reduce the properties of extraposition to an extended version of PUB. He argues that adjunction to the right is a fourth type of position - as we will see more clearly below. However, I think that most of his necessary prerequisites in order for this analysis to work rely on problematic assumptions.

First notice that this account crucially depends on the assumption that extraposition is derived via movement. All the arguments against a movement approach for extraposition discussed in section 1. apply. Moreover I do not agree with his line of reasoning for the necessity of movement especially to derive the clause boundedness of extraposition. Discussing the analysis of C&R, Müller states:

> *"Under this assumption the apparent NP-violations are straightforwardly accounted for - there is no movement in the first place. However, as concerns the clause-boundedness of extraposition, one has to resort to some additional condition. C&R (1990, 26&41), for instance stipulate that the NP and the extraposed item must govern each other [...]. Ideally, though, the clause-boundedness of extraposition should follow without additional assumptions from independently attested principles of movement".*
> (Müller 1994: 12f.)

There are several problems with this line of reasoning. First, if there are two properties X and Y, where X apparently suggests an analysis A and Y suggests an analysis B it cannot simply be concluded that one of the properties A or B respectively has priority over the other one. Secondly, it is not clear at all why it should "ideally" be the case that if a relation is clause bounded it should indicate that this is a relation derived via movement. Notice that there are instances of movement which are NOT clause bound. On the other hand there are base-generated relations that are clause bound (as for example binding of anaphora). Thus if a relation is clause bound it cannot be concluded that it should be derived by a movement analysis. Thirdly, it is exactly this

clause boundedness which does not really fall out from the PUB. It is true that under this account 'standard-successive-cyclic' movement is not an option. In that case the variable would first be bound by the trace in SpecCP and then it would be bound by the extraposed constituent which is right-adjoined. Thus the original trace would be ambiguously bound by two elements in two different types of positions. However, Müller has to stipulate that (although right-adjunction to NP and IP is allowed) right-adjunction to CP is not an option. Müller admits that there is no straightforward reason for this property (which would fall out from the PUB). The only (empirical) evidence for this stipulation is that apparently overt right-adjunction to CP is also not possible. However, as we have seen there is a counter-example to this generalization, i.e. the cases where upward-boundedness is apparently violated. The relevant example is repeated here for convenience:

(84) *$[_S$Which spy t does e$_j$ $[$does Angelton believe $[_S$ e$_j$ that $[_S$ Burgess recruited e$_j]]$ who ultimately became a mole$_j]]$* (G&M 1984,: 16)

Since Müller assumes that the extraposed constituent has to antecedent-govern its trace he would have to assume that adjunction to CP is an option. His stipulation does not even receive full empirical support.

Fourth, Müller cannot account for the Nesting Effect in a straight-forward way. He assumes extraposed constituents to invariantly adjoin to IP - and this is again a rather constructions specific statement. I conclude that even the first argument in favor of his analysis is not a legitimate argument. Clause-boundedness does not necessarily imply movement - i.e. to the contrary it suggests base-generation. In order for the analysis to really work he has to add a stipulation, which is neither theoretically nor empirically justified. Another crucial assumption for his analysis to work is the following:

> *"The point that will turn out to be important , however, is that wh-movement, scrambling, and topicalization are all movement operations that are* formally different from extraposition, *in the sense that the landing site of extraposition is not the same as the landing site of any of the other movement types. And indeed, given that extraposition is* rightward *movement, whereas wh-movement, scrambling, and topicalization are all instances of* leftward *movement, this follows without further stipulation. [...] as a consequence the landing site of extraposition is formally different from the landing site of any leftward movement type."* (Müller 1994: 16)

I do not think that this point is legitimate either. I have already discussed the interaction of linear ordering and hierarchical structure in previous sections. There I have concluded that within hierarchical structure left- and rightness should not play a role at all. I take reference to rightness a construction specific statement. However, the claim that right-adjunction is formally different to left-adjunction is not the only part in Müller's analysis where he has to explicitly refer to rightness. In order for his proposal to work he has to stipulate that right-adjunction (to NP and IP) is an option (in German) whereas left-adjunction is not. In order for PUB to derive the correct result he has to first stipulate that right-adjunction is a formally different landing site and second he has to stipulate that elements (which can be extraposed) can only right-adjoin but not left-adjoin. Moreover, the main goal, namely to derive the properties of extraposition from independent principles is undermined, in that there are purely construction-specific statements involved.

Müller however argues that the empirical facts support the view that for certain constituents it is only possible to right-adjoin but not to left-adjoin. (Notice that the precedence constraint derives this empirical generalization). This part of the analysis accounts for apparent island-violations. Müller assumes that NPs in general are barriers for movement. One possibility to circumvent this barrierhood is taken to be the possibility of adjunction to this barrier (in the sense of Chomsky 1986). If adjunction to the right is a possible option then right-ward movement is predicted to be less restricted (i.e. it can cross apparent barriers)[29].

Besides these main theoretical problems there are also some empirical problems which I will briefly discuss. First, it seems to me that if the properties of extraposition are really to be derived by the PUB then it is predicted that wh-islands should still make a difference in judgements. First consider 'long-distance-scrambling':

(85) (a) *?Ich habe einen Pudding gesagt daß ich esse.*
I have a jelly said that I eat
(b) **Ich habe einen Pudding gesagt, wann ich esse.*
I have a jelly said, when I eat

There is a clear contrast between the sentences in (85a) and (85b), the latter being degraded. (85) could suggest that a PUB violation is less strong than a wh-island violation. In (85a) long-distant scrambling can apply successive cyclically violating only the PUB. In (85b) no such option is available. The sentence can be argued to be

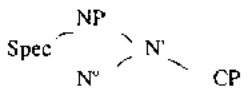
[29] Notice that it is crucial here that right-adjunction to CP is not an option, as discussed above.

a wh-island violation. The sentence is degraded. Thus long distance scrambling without the intermediate step via SpecCP is worse than the PUB-violation. This seems to be an empirically forced result. However, then there should also be a contrast between extraposition violating upward boundedness depending on whether an element is already occupying SpecCP (the intermediate landing site) and extraposition violating only PUB. This is not a correct prediction:

(86) (a) **Antje hat immer zugegeben daß der Versuch scheitern muß wenn sie guter Laune war, mit vier Bällen zu jonglieren.*
Antje has always admitted that the attempt fail must when she good mood was, with four balls to balance
(b) **Antje hat gefragt, wann der Versuch scheiterte, als sie guter Laune war, mit vier Bällen zu jonglieren*
Antje has always admitted when the attempt fail must when she good mood was, with four balls to balance

Both sentences are equally ill-formed no matter whether the intermediate landing site is available or not.

Another problem for Müller's analysis is that he necessarily has to assume that relative clauses are generated in a position dominated by N'. Otherwise the NP would not count as a barrier and he couldn't derive the left-right asymmetries. An NP construed with a relative would have a D-S representation like below:

(87)


Under such an approach the determiner and the NP would not create a constituent without the CP. It seems however that under such an approach similar arguments as against the Raising (right-branching) analysis for relative clauses obtain. Those arguments will be discussed in Appendix II. A similar problem arises because Müller has to assume that CPs can be in complement position of *es* in German. It is however in general impossible for a pronoun, to have a complement. Moreover since Müller assumes possible S-structures to be support for covert representations he has to face the problem for why it is not so good to have the CP in adjacent position to this pronoun. The same obtains to ProPPs. However here I will not go into a detailed

discussion of an appropriate analysis for ProPP[30].

There is another problem for Müller. Remember the paradigms PCI can explain (cf. section VI.1.): 1) no left-adjunction of the identifier, 2) no movement of the identifier to a position preceding the identifyee 3) no movement of a constituent containing the identifier to a position preceding the identifyee (and 4) no movement out of the identifier). Müller's constraints involving reference to rightness can only account for the first two properties. Neither the PUB nor the prohibition of left-adjunction can account for the third property. Consider again an example where the extraposed constituent is adjoined to a topicalized VP the antecedent being left behind in the matrix.

(88) *[geglaubt [daß Maria Bier trinkt]$_i$]$_j$, hat Peter (*daran$_i$/*es$_i$) nicht t$_j$*
[believed [that Mary beer drinks]$_i$]$_j$, has Peter (*thereon$_i$/*it$_i$) not.

PUB cannot rule out (88). There is no ambiguous binding since there is no subsequent movement of the 'extraposed element' itself, rather it moves along contained in a different constituent.

4. On some properties of the correlate

The last sections were dealing with the properties and restriction of the relation under consideration. In this section I want to now address the question of the nature and the restriction of the correlate (i.e. the identifyee) and the various accounts for the properties that are to be detected.

We have already seen some of the accounts given for the problem of the restriction imposed on the identifyee. I will thus briefly repeat the main properties of these accounts and turn towards a criticism of them. There are in principle two different possibilities to account for the properties: a structural one (which as far as I can see is restricted to a movement approach). Under such a view the question to be addressed is whether or not XP allows for movement out of it, resp. whether a trace is licensed within XP. (For such an account I will again discuss Guerón (1980). The other approach is more or less an interpretational account where the question arises as to whether XP itself can be related to another constituent in the domain established in previous sections, i.e. whether XP licenses YP in a position which is not adjacent (to

[30] cf. Wiltschko (1993) for an analysis.

choose a rather neutral term). This approach is possible in a movement as well as in a base-generation approach. The restriction on XP can be empirically described as the following well known property of extraposition:

(89) A nonrestrictive relative cannot be extraposed.[31]

It might seem a bit strange to describe the property of the identifyee by making reference to a property of the identifier or rather a property of the relation between the identifyee and the identifier respectively. Notice however, that the interpretation of restrictive vs. nonrestrictive modification is also (or maybe totally) a function of the interpretation of the identifyee. Thus we can state the claim above differently, i.e. just referring to a property of the identifyee itself:

(90) A fully referring/denoting expression does not license extraposition.

Guerón's (1980) Name Constraint. As we have already discussed, Guerón's analysis is within a movement approach towards extraposition. Her account for the generalization in (90) is captured by the so called *Name constraint*:

(91) *The Name Constraint* (first version)
A complete referring expression cannot contain a variable which is not bound within it.
The Name Constraint (second version)
A Name may not contain an empty argument position (Guerón 1980: 666)

For the Name Constraint to exclude extraposition out of Names Guerón has to assume that relative clauses and other elements 'classified' as identifiers are base-generated within the NP where it already counts as a complement. However, if a complement can be any constituent governed by the complement-taker then it is not obvious that there is really a reason to assume this constituent as base-generated within the NP in the first place. Notice moreover that variables are defined as follows:

[31] Of course it is not only relative clauses that constitute our empirical domain - thus we should extend this claim to saying that a non-restrictive modifier cannot be extraposed. However as we will see the exact formulation of this restriction does not really matter since it is rather a restriction on the element to which the extraposed element is related.

(92) Variable $=_{def}$ $[_{NP}e]$ in A-position locally A'-bound and operator bound.[32]
(Cinque, 1991)

This means that the complement must be base-generated within an A-position, i.e. it counts as an argument. It seems to me that the definition of the Name-Constraint implies that being a complement implies being an argument. I think that this undermines the necessity of the complement principle, i.e. introducing the complement relation as a primitive relation. Either the definition of the Name Constraint (and with it the definition of variables) is reconsidered or the Name Constraint cannot exclude cases where a non-argument (that should still be a complement in the sense of the complement-principle) is extracted. So let us suppose that this inherent problem can be dispensed with. However there are some more problems to be detected within Guerón's account which are not as general as the one just discussed but which are particular to her analysis. Consider Guerón's assumptions concerning the interpretation of NPs:

(93) *NP Interpretation:*

A Name is a complete referring expression. It designates a unique object or individual (or set of these) in the world of the discourse, either directly through the use of proper names or deictic expressions (John, that man), or indirectly, by means of complements containing direct referring expressions (the girl who sits next to you, some of those books)

A quantifying phrase is an operator ranging over a set of entities. It does not designate a unique individual.

I have already discussed some theoretical problems with this account w.r.t. the assumption of traces. The spirit of this account is intuitively plausible: it seems to be

[32]At this point let me briefly mention that there is an inherent problem in the last formulation of the Name Constraint which is however a problem concerning the definition of a variable in general. Notice that this sort of definition is equivalent to the first definition she gives under the assumption that the definition of variables is as follows:

(i) Variable $=_{def}$ $[_{NP}e]$ in A-position, locally A'-bound by an operator
(cf. Chomsky 1981)

The general problem at stake here is arising from the definition of variables on the one hand and the prohibition against vacuous quantification on the other hand. Of course it is perfectly grammatical to wh-move adjuncts. So we end up with a contradiction: wh-words are inherent operators and operators have to bind a variable but variables are defined as occupying an A-position. However the trace of adjuncts (the variable we would need) does not occupy an A-position (at least under standard assumptions). The most appropriate way for a solution seems to be a rethinking of the definition of A- vs. A'-positions, which seems to be necessary anyway (cf. Chomsky (1992)).

an inherent contradiction to be a full referring expression and at the same time containing a variable - which needs to be bound from outside. However it is not really clear why this really provides us with an explanation for the facts: containing a variable could also render an otherwise complete referring expression into a non-complete-referring expression, without necessarily resulting in ungrammaticality.[33] Moreover, there seems to be a contradictory claim in the definition Guerón gives which has to do with the assumption concerning the possibility to indirectly designate a unique object or individual by means of complements containing direct referring expressions. As we have seen definite NPs can be used to introduce a new discourse referent by means of an identifier as in the example below (repeated from chapter IV, for convenience):

(94) Q: *Weißt du ob Maria ein Buch gelesen hat?*
'Do you know whether Mary has read a book? '
A1: *Ja, Maria hat ein/*das Buch gelesen.*
'Yes Mary has read a/*the book.'
A2: *Ja, Maria hat ein/das Buch gelesen, das Hans ihr geschenkt hat*
'Yes, Mary has read a/the book John gave to her.'

In this case the identifier does exactly what Guerón means with indirectly designating a unique object. At the level where NP-interpretation takes place (Guerón's second LF-level, which is intended to do similar things as domain D used in the approach defended here) the NP under consideration is referring to a unique object, by virtue of the identifier, i.e. it would fall under the definition of 'Names' in the sense of Guerón. However, as we can see above, extraposition is still possible - although it should violate the Name Constraint.[34]

In addition I have argued (following proposals by Bennis 1987, Cardinaletti 1990) that in case of *'es'* or a ProPP serving as the correlate within the matrix we find the same situation. In these cases Guerón is confronted with a severe problem: she would have to say that the extraposed element is moved out from a position which is

[33] Notice that Guerón's analysis is somehow encoding an interpretive account into a apparently structural account, i.e. it is assumed that within a given domain (which is defined interpretive) a trace is not licensed (because of some interpretational properties of this domain, i.e. the complete referring expression).

[34] Notice that Guerón is just dealing with instances of PP-extraposition and the examples given here as counter-evidence involve relative clause -(or 'argument' clause) extraposition. The Name Constraint is formulated generally enough that it should equally apply to these instances of extraposition.

somehow adjacent to the pronoun (the CP being a complement to the pronoun). However, pronouns presumably fall under the definition of Names (within Guerón's assumption) thus extraposition should be ungrammatical - violating the Name constraint. In this case Guerón could still argue that the Name Constraint does not apply since pronouns do not count as Names. However, as we have seen there are also deictic pronouns that can serve as the correlate in the matrix clause as for example in (95) below:

(95) *Ich habe das geglaubt, daß Maria gerne Bier trinkt.*
I have d-pron. believed, that Mary likes beer drinking.

Moreover we have seen that it is a property of (certain) cataphoric elements that they behave similar to deictic expressions.

But now let me turn to the data Guerón discusses as evidence for the adequacy of the Name Constraint. She starts with the following minimal pair:

(96) (a) *A book was published about linguistics.*
(b) **That book was published about linguistics.*

At first sight it might seem that the relevant generalization is the definiteness of the determiner of NP from which extraposition takes place. However, this generalization is not even empirically adequate as Guerón notices:

(97) (a) **A particular book came out by Chomsky.*
(b) **A certain book came out by Chomsky.*
(c) **A book we had been looking forward to came out by Bill.*
(d) **An article in the Times was circulated on linguistics.*

Guerón continues her argumentation in showing that the definiteness vs. indefiniteness of the determiner is not enough for determining the referential properties of a given NP. She claims that indefinites can differ as to whether they are interpreted as being specific or nonspecific whereas definite NPs can differ w.r.t. being referential vs. non-referential (attributively used to use more traditional terms). Consider the following paradigm, especially the intended interpretation of the underlined NP:

(98) (a) *Mary wants to marry a Swede* (ambiguous)
(b) *She met him in Paris* (specific)
(c) *...although she's never met any* (nonspecific)

(99) (a) *I would like to read the review of John's book* (ambiguous)
(b) *...that's lying on the table over there* (referential)
(c) *...if you ever write it* (nonreferential)

Following Partee (1972) Guerón assumes:

> "that the specific-nonspecific ambiguity of indefinite NPs should be treated on a par with the referential-attributive ambiguity of definite NPs discerned by Donnellan (1966). Adopting this suggestion, May (1977) suggests that specific indefinites and [referential][35] definites be considered *Names*, in the Fregean sense: complete referring expressions. Nonspecific indefinites and attributive definites on the other hand, would be treated as *Quantifying Phrases*. This distinction is crucial at the level of logical form." (Guerón, 1980: 666)

Assuming this Guerón gets the result that wh-movement (leftward movement) on the one hand as well as extraposition (rightward movement in her terms) are restricted by the same constraint (i.e. the Name Constraint). The following data suggest that this prediction is indeed empirically confirmed. It seems to be motivated that the movement approach is on the right track. The following facts involving wh-movement should be captured by the same generalization that is responsible for the possibility or impossibility to extrapose:

(100) (a) *What subject did they publish a book about?*
(b) **What subject did they publish that book about?*
(c) **What subject did they publish a certain book about?*
(d) **What subject did they publish a book you had been looking forward to about?*

The LF-representation for wh-movement, QR (movement at LF) and extraposition is given below to show how the Name Constraint derives these facts:

wh-movement:
(101) (a) *Who did you see a picture of?*
(b) **Who did you see that picture of?*
(c) LF: For which x, you saw ($_{NP}$ that picture of x)

[35] in the text it says 'attributive' definites. However, this is clearly a typo.

QR:

(102) (a) *John took a picture of everyone* (where *everyone* has wide scope)
(b) **John took that picture of everyone*
(c) LF: For all x, John took ($_{NP}$ that picture of x)

Extraposition:

(103) (a) *Books were sold on linguistic theory*
(b) **Those books were sold on linguistic theory*
(c) LF: ($_{S}$($_{S}$ were sold, ($_{S}$($_{NP}$ those books $[_{PP_j}$ e])($_{VP}$ ...V_i ...)$_{S}$)($_{PP_j}$ on linguistic theory)$_{S}$))

In all these examples we get an LF-representation where a variable is free within the Name. Moreover, the Name Constraint also captures the following facts where a pronoun acting as a bound variable is just wellformed if it is not contained in a Name rather it has to be within an attributive NP - it is again relevant that the definite determiner does not necessarily give rise to a Name (i.e. referential) interpretation:

(104) (a) *Any dog loves the person who feeds it*
(b) **Any dog loves Bill, the person who feeds it*
(c) **Any dog loves the fellow you introduced me to, who feeds it.*

The sentence in (104a) is wellformed although the object NP (*the person who feeds it*) contains a variable bound by the quantifying expression outside it - however this NP is considered to be attributive rather than referential. On the other hand the Name in (104b) and the deictic expressions in (104c) force a referential interpretation of the NP resulting in a non-wellformed representation at LF. It seems to me that this paradigm is not so straightforward as Guerón claims it to be. First of all it is not clear why the relative pronoun (being an operator) does not intervene in the binding relation between the quantifying expression and the pronoun (that is to be interpreted as a bound variable). This is however true in all the examples in (104). So it is not clear to me how the contrast really arises. The 'variable' interpretation could also be dependent on the interpretation of the head NP as in the example below discussed in Williams (1994):

(105) *Every hospital administrator hates the hospital*

In (105) the interpretation of *the hospital* is dependent on *every hospital administrator*. This fact is argued by Williams to be crucial evidence for the assumption that dependency and coreference are to be analyzed as different relations (governed by different constraint). If true, then *it* in the examples above can be analyzed as being dependent (without being coreferent) on the interpretation of the head NP. Since the

reference of this NP varies dependent on the quantifying expression *any dog* the reference of *it* can also vary. Coreference is however an independent property. In case of the non-restrictive relative clauses, the reference of the head NP does not depend on the quantifying expression. Therefore the pronoun cannot be either. It would have to be anaphorically related to the quantifying expression. In this case it is however not a interpreted as a variable anymore. It would not violate the Name Constraint. Notice however that it is in general not possible to anaphorically relate a pronoun to a quantifying expression as exemplified below:

(106) *Any dog loves John. *He feeds it well enough*

This suggests that the sentences in (104) are independently ruled out. In addition the Name Constraint has severe problems in deriving the contrasts detected there.

The last fact Guerón discusses as evidence for the Name Constraint is the ambiguity of the NP in (107a) below, as opposed to the non-ambiguity of the same NP if the modifying relative clause occurs in extraposed position:

(107) (a) *Those people whom you want may come.*
(b) *Those people may come whom you want.*

The distinction between Names and Quantifying Phrases - Guerón continues to argue - must be stated in the Grammar independently of extraction operations (implying that the Name Constraint is not a construction specific constraint). She mentions some facts where the distinction becomes relevant:

(108) (a) Properties may only be predicated of Names. [16]
(b) Backwards coreference is possible only with Names.
(c) Certain syntactic frames allow only Quantifying Phrases (i.e. *there* Insertion in English)

At first sight the fact that leftward movement seems to be constraint by the same restriction as extraposition seems to be evidence for the movement nature of

[16] This generalization seems to be problematic w.r.t. the following paradigm:
(i) *Any man with a hat is nice*
(ii) *That man is nice.*
Especially on basis of the previous argumentation concerning extraposition and the impossibility to have a bound variable within a Name this paradigm does not follow straightforwardly. (Thanks to Edwin Williams (p.c.) for pointing this problem out).

extraposition. However, I have already rejected the movement analysis. Therefore we loose the explanation which would generalize two similar phenomena. What we have to do is showing that both facts are just similar at surface. The rejection of this claim would be even more convincing if we would find data where the generalization does not hold anymore, i.e. for supporting our claim we have to find examples where wh-extraction (or leftward movement in general) is possible, but extraposition is not. A second piece of evidence would be examples where we find the reverse situation: extraposition being possible but leftward-movement leading to ungrammaticality.[37]

Concerning the latter case we have already seen that leftward-movement always leads to ungrammaticality if the moved element is coindexed with an element (i.e. the identifyee in our terms) in the matrix whereas extraposition results in a wellformed configuration.

An instance of the other potential counter-evidence to Guerón's claim is the fact that preposition stranding in English is allowed in case of left-ward movement but is disallowed in case of extraposition (i.e. rightward movement):

(109) (a) *Who did Peter talk [$_{PP}$ about t].*
(b) **Peter talked [$_{PP}$ about t] yesterday Mary.*

Notice that Guerón assumes that:

"the PP trace [...] has the same status as a variable on the level of SI-2"
(Guerón, 1980: 666)

This seems to suggest that PPs and NPs are equated. So PPs presumably count as Names. Assuming this Guerón can not correctly rule out (109b) but at the same time it would be predicted that (109a) is equally ruled out. If it is not assumed that the Name Constraint applies to extraction out of PPs, then Guerón still has a problem with the asymmetry of left- vs. rightward movement in case of P-stranding. There is no straightforward reason as to why (109b) should be ill-formed if one assumes a movement approach for extraposition - both (109a) and (109b) involve an A'-chain. In a basegeneration approach the ill-formedness of (109b) could be reduced to an instance of a violation of the case filter. If a trace is assumed to be left by extraposition we find

[37] Notice that Guerón shows that the distinction between Names and Quantifying phrases is relevant for the well-formedness w.r.t. ill-formedness of the element in certain other syntactic environments not involving movement (i.e. the environments mentioned in (108a-c). Therefore the evidence for the movement approach is already undermined.

a chain. The foot of the chain sits in case position and therefore the case filter is satisfied. I take this as another evidence of assuming extraposition to be a base-generated phenomenon.[38]

Another instance where the Name Constraint is apparently violated is the phenomenon known as Split Topicalization. (Of course the following argument against the Name Constraint only applies under the assumption that Split Topicalization is an instance of move alpha[39]. Consider the following German example:

(110) *[Schuhe]$_i$ habe ich heute [die braunen t$_i$] an*
shoes have I today the brown on
'As for shoes, today I am wearing the brown ones'

The NP falls under the definition of Names, however it contains a trace (which is presumably a variable). However the extraction is wellformed - contradicting the Name Constraint.

Notice that the analysis suggested here captures the intuitive plausible assumption of Guerón, namely that fully referring expressions cannot contain an open position, by means of the following assumptions. It has been argued that a DP can induce the need of further licensing if it does not have enough descriptive content to pick out the relevant discourse referent. The identifier provides further descriptive content in order to pick out the intended discourse referent. Therefore it is expected that fully referring expressions, i.e. expressions that have enough descriptive content to pick out the intended referent, do not induce the need for identification. Therefore it follows that adding further modification (e.g. a relative clause) is not an instance of Identification. A non-restrictive relative clause does not provide content to identify the discourse referent, rather it adds further information about the discourse referent. Therefore we expect different licensing conditions. Although I have not been addressing the issue of non-restrictive modification it is at least expected that they behave differently. I will leave the exact analysis of non-restrictive modification (in the light of the present proposal) as a matter of future research.

[38] Of course assuming this we would have to say something about Heavy NP shift - a phenomenon I do not address in this dissertation.

[39] cf. Haiden 1994 for a discussion of the problems with this assumption.

5. Conclusion

The discussion in this Appendix was intended to show how the analysis presented in this dissertation differs from previous analyses of extraposition. The main and crucial difference is that it can derive a stipulation which is found in all previous analyses. This concerns the obligatory rightness of extraposition. The second difference to most of the traditional analyses is that it does not have to mention a specified adjunction site of extraposition.

APPENDIX II.
Against a rightbranching analysis for extraposition.

0. Introduction

One of the main goal of the present dissertation is to give an adequate analysis of extraposition. We have seen some theoretical and also empirical reasons to assume right-adjunction for extraposed constituents. In recent developments there are some debates as to whether right-adjunction (or adjunction in general) is an option for UG. Therefore an analysis which heavily makes use of the possibility to right-adjoin constituents should at least address this issue more carefully. I will give a brief overview over the 'right-branching'-proposals found in the literature. Then I will compare the 'right-branching' analysis with the (traditional) right-adjunction analysis. I will present some empirical facts that support the latter approach.[1] Finally, I will briefly sketch how the analysis presented here could be interpreted in accordance with the main claim in Kayne (1993) that leads him to assume a strictly right-branching analysis in the first place.

1. The right-branching analysis (Kayne 1993, Zwart 1992, Haider 1993)

In recent developments within syntactic theory there are three main proponents of the right-branching analysis. Kayne (1993) argues for a strictly right-branching phrase-structure on basis of considerations having to do with linearization. He argues that hierarchical structure unambiguously determines linear ordering. The main claim is that asymmetric c-command maps into linear precedence. Therefore it is predicted that movement and also base-generated adjunction to the right is not an option for UG. Right-adjunction would result in an output where an element X asymmetrically c-commanding an element Y does not precede Y but rather X follows Y. The second prediction is that phrase-structure is universally head-initial. Zwart (1992) applies the idea that X'-structure is universally head-initial. He argues that this is also true for Dutch (and German) VPs, which traditionally have been analyzed as being head-final. Elements to the right of the verb are assumed to occupy their base-position, i.e. the complement position of the verb. Zwart's main empirical arguments

[1] Cf. also Büring & Hartmann (1994) for empirical arguments for the right-adjunction analysis as opposed to the right-branching analysis.

for the assumption that Dutch and German have head-initial VPs are as follows:

- If complements are generated as a sister to the right of the head it is expected that certain constituents can occur in this position. He claims that so called 'extraposed' CPs and PPs are really occupying their base-position. The reason that NPs never occur in this position is that they have to move to the Specifier of a functional projection in order to receive case under Spec Head Agreement. CPs and PPs do not need case and therefore they can stay in situ.
- Complement CPs in 'extraposed' position do not induce an island for extraction. Therefore it seems reasonable to assume that they occupy an argument position.

Furthermore he assumes that the verb moves to the head of a projection immediately dominating the VP; he calls this projection PredP (=Predicate Phrase) which I will discuss later on.

Notice that the first argument is also valid within a right-adjunction analysis. It can be assumed that extraposition is only possible for elements that do not need case. Therefore it can be assumed that at S-structure there has to be a strict local relation between the case-assigner and -assignee at S-structure. Zwart's second argument is similar to the one in Bennis (1986) and Cardinaletti (1990) discussed in section VI.1.4.1. There it has been shown that even under a head-initial approach one has to assume the possibility for extraposition. Extractability was shown to not being influenced by extraposition. The evidence has to do with the fact that an intervening (preceding) extraposed clause does not render extraction ungrammatical.

Haider (1993) explicitly argues against extraposition as right-adjunction. He also assumes that phrase-structure is strictly rightbranching. The difference between him and Kayne (resp. Zwart) is however that Haider still assumes a head-final VP for German. In his approach the serialization [V CP] results from movement of the verb to a position preceding the 'extraposed' element. For doing so he assumes VP-shells in the sense of Larson (1988).

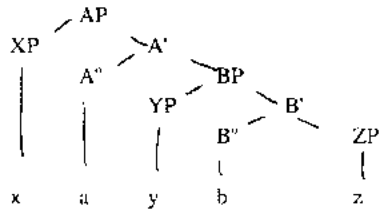
Most of the arguments presented below hold against both approaches and if it is not necessary to distinguish them I will refer to all of them as the right-branching analysis. Notice that all the arguments presented in this appendix are evidence for the necessity of right-adjunction. They are not necessarily arguments against a head-initial VP. It will simply be shown that even within a head-initial approach extraposition as right-adjunction has to be a possible option. As was shown

in Appendix I, a movement approach to extraposition has a lot of problems on its own. Therefore it follows that right-adjunction must be possible in the base. Moreover I will also present some arguments indicating that the head-initial analysis for German VPs has some problems of its own.

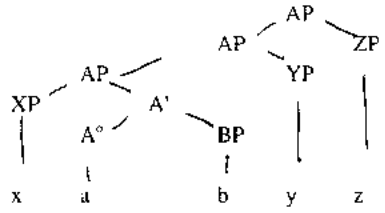
2. Rightbranching vs. Right-adjunction

It was implicitly shown in the previous chapters that there are several empirical arguments for a right-adjunction analysis of extraposition. Before I will go into detail I will summarize the difference between the right-branching and the right-adjunction analysis. First compare the two different phrase-structures resulting from either approach:

(1)



(2)



It is clear that the relation between dominance and precedence is mirrored in these structures: In a right-branching tree (1) the linearly first element is hierarchically the highest. The latter an element occurs the lower it is. Right-adjunction mirrors this situation (2). The element that is dominated by the lowest node linearly precedes every other element: the higher an element is the latter it occurs within the linear order. Of course this crucial difference is expected to have an effect in various

syntactic environments.[2] The difference in hierarchical structure has some consequences concerning constituenthood. In a right-branching analysis an element that follows all other elements in linear ordering is expected to be within the lowest maximal projection. Therefore it is expected that it forms a constituent with the lowest head. In the structure in (1) above ZP should form a constituent with B° and YP, namely BP. However, under the right-adjunction analysis (as represented in (2)) YP can form a constituent with AP but it need not. Adjunction creates a configuration where this option arises. In addition it is expected that ZP can also form a constituent with AP. However, if it forms a constituent with AP then by transitivity of the dominance relation it is necessarily the case that all constituents linearly occurring between A° and ZP (i.e. all elements adjoined to AP below ZP) have to be part of this constituent as well. ZP is part of the highest projection of AP.

Given this situation we can apply the traditional tests for constituency to decide between the two approaches: these tests concern 1) movement, 2) deletion and pronominalization and 3) coordination. As we will see all these tests favor a right-adjunction analysis for extraposition.

2.1. VP-topicalization

Let us first turn to the movement-test. In a right-branching analysis (like (1)) the following situation is expected: If BP is moved then ZP has to be moved along with BP since it necessarily forms a constituent. On the other hand under the right-adjunction analysis (as represented in (2)) we expect a different behavior: If AP is moved, then YP is expected to be able to either move along with AP or alternatively to stay behind in its base-position. What we find is exactly the latter behavior indicating that the right-adjunction analysis is on the right track.

A crucial empirical problem for the base-generation analysis involves VP-topicalization. As the paradigm below shows it is either possible to move a complement CP along with the topicalized VP or to leave it behind (in extraposed position):

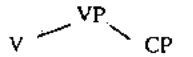
[2] Notice that we have already seen a crucial evidence for this mirroring effect in section VI.3. There I presented data where the scope of adverbials suggested that hierarchical structure is indeed symmetrical.

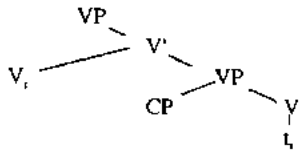
(3) (a) *[Der Maria gesagt]$_{VP}$ hat Hans. [daß Peter ein Spion ist]$_{CP}$*
[the Mary said] has John [that Peter a spy is]
'...and said to Mary has John that Peter is a spy'

(b) *[[Der Maria gesagt]$_{VP}$ [daß Peter ein Spion ist]$_{CP}$]$_{VP}$, hat Hans.*
The Mary said, that Peter a spy is, has John

The sentence in (3a) receives a straightforward explanation under the right-adjunction analysis. Under the right-branching analysis things do not follow quite so simply[3]. Assuming that the latter an element occurs the deeper it is embedded one has to assume that the equivalent sentence without topicalization has the following structure:

(4) (a) *weil Hans der Maria gesagt hat, daß Peter ein Spion ist*
since John the Mary said has, that Peter a spy is

(b)


(c)


Under the assumption that the German VP is head-initial the structure in (4b) has to be assumed for to the sentence in (4a). Under Haider's analysis, assuming a head-final VP, (4c) would be the corresponding phrase-structure for (4a). The verb is moved to a position preceding the complement CP which itself stays in its base-position. However, if the VP is topicalized one would expect that the complement CP obligatorily has to move along with the VP. Although this is also a possible derivation (as (3b) indicates) it is not the only one available (as the sentence in (3a) shows). It is evident that under a right-adjunction analysis for extraposition there is no problem with the paradigm in (3) since the complement CP is adjoined to VP and topicalization can move the higher or the lower VP node as expected. However, (3a) creates a serious problem for the right-adjunction analysis. There is no obvious way to move the VP stranding behind the complement CP that is contained within this VP.

A possible solution for the right-branching analysis would be to assume that (like NP complements) the complement CP is moved out of the VP into the

[3] Culicover & Rochemont (1994) discuss some possibilities to derive some of the sentences in a right-branching analysis. Here I will not present these possibilities.

Specifier of a functional projection. However, if this would be possible we would expect this sentence to have the same status (w.r.t. interpretation as well as well-formedness) as the equivalent sentence where the VP stays in situ and the complement CP occupies a position preceding the verb. As the following example shows there is a difference between extraposed CPs and CPs that occur in a position preceding the verb (i.e. in the 'Mittelfeld'):

(5) (a) *??Hans hat [daß Peter ein Spion ist]$_{CP}$ der Maria gesagt*
John has, that Peter a spy is, the Mary said
(b) *?Hans hat der Maria [daß Peter ein Spion]$_{CP}$ ist gesagt*
John has the Mary that Peter a spy is said

There is a clear contrast between (4), which is fully grammatical and the sentences in (5), which are degraded (especially when the complement CP occurs in a position preceding the indirect object (5a)). Moreover in case the CP is not extraposed then the interpretation is different (as discussed in section VI.4.)[4]. The same argument can also be used for an analysis that does not make use of head-initial VPs like it is assumed in Haider (1993). Again there is no possible derivation for the sentence in (3a).[5]

The same problem arises in case of extraposed relative clauses (which are not analyzed in Zwart (1992) but in Haider (1993) and in Kayne (1993)). Kayne (1993) makes use of an analysis for relative clauses that is in the spirit of Vergnaud (1974). It is suggested that extraposition of relative clauses has to be rethought much like

[4] cf. Trinker (in preparation) for an elaborate analysis for CPs occurring within the German 'Mittelfeld'.

[5] Haider mentions the fact that Left-dislocation of the VP is not possible when the complement CP stays in its final position:
(i) **[Gesagt]$_i$ das hat der Mann keinem freiwillig, wo ich wohne$_i$*
said, d-pron. has the man nobody voluntarily, where I live
*[Das Haus gezeigt] das hat der Mann keinem freiwillig, in dem ich wohne
the house shown d-pron. has the man nobody voluntarily, in which I live
(Haider 1993: 8a/b)
Haider uses these examples as evidence for the rightbranching analysis since all these sentences are grammatical if the complement CP is moved along with the VP:
(ii) *[Gesagt, wo ich wohne]$_i$ (das) hat der Mann keinem freiwillig$_i$*
[Das Haus gezeigt, in dem ich wohne] (das) hat der Mann keinem freiwillig.
His argument is correct for Left Dislocation but it does not hold for VP topicalization since the sentences in (3) are grammatical. I do not have an analysis for the paradigm in (i). The marginality of these sentences might have to do with the fact that the left-dislocated element behaves more like a discourse antecedent (as discussed in chapter VII.).

the process of *quantifier floating* was reanalyzed as *quantifier stranding*. Therefore he has to assume that the head of the relative is base-generated within the relative clause[6] and subsequently moves out of it. The relative pronoun is analyzed as a copy of the trace. Therefore extraposed relative clauses are in base-generated position (i.e. within the complement position of the verb). The head-NP moves to a position preceding the verb (i.e. into the specifier of a functional projection) in order to receive case.

Again the problem concerning extraposition and VP-topicalization is as follows: If a VP containing an object NP that is the head of a relative clause the relative clause can be left stranded behind:

(6) (a) *[[Den Mann]$_i$ gesehen [der am Fest rauchte]$_i$]$_{VP}$ hat Maria nicht*
The man seen who at the party smoked has Mary not
'...and seen the man who smoked at the party Mary has not'
(b) *[Den Mann]$_i$ gesehen]$_{VP}$ hat Maria nicht [der am Fest rauchte]$_i$.*
the man seen has Mary not who at the party smoked

Again under the right-adjunction analysis one would have to assume that the relative clause that occurs in sentence-final position (6b) is stranded behind in a specifier position of a functional projection. This possibility can easily be rejected. Relative clauses can never occur in such a position if the head NP is topicalized on its own:

(7) (a) **[Das Mädchen]$_i$ hat niemand [das gerne raucht]$_i$ gesehen*
the girl has nobody who likes to smoke seen
(b) *[[Das Mädchen]$_i$ gesehen]$_{VP}$ hat niemand [das gerne raucht]$_i$.*
the girl seen has nobody who likes to smoke

The right-adjunction analysis derives these data in a straightforward way.

2.2. More deeply embedded antecedents

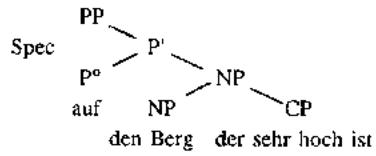
There are several different constructions that constitute the same problem for the stranding analysis of extraposed relatives. All of these data involve an antecedent which is somehow more deeply embedded. Again the empirical data favor a right-

[6] Vergnaud (1974) argues for such an analysis mainly because of the possibility to have a relative clauses within an idiom chunk. This suggests that the relevant NP can only be generated within the relative where the verb of the idiom occurs.

adjunction analysis. The simplest case are relative clauses with a head NP embedded in a PP:

(8) *Peter ist [auf [einen Berg]$_i$]$_{PP}$ gestiegen, [der sehr hoch ist]$_i$.*
Peter is on a mountain climbed which very high is

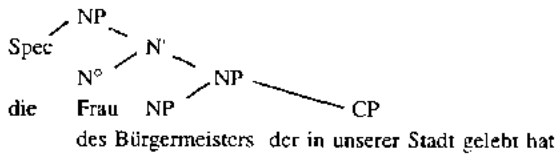
Under a stranding analysis one has to assume that it is just the NP that was moved out of the relative clause. The relativized constituent is only an NP (*der Berg*). It can be assumed that the NP adjoins to the relative clause forming an NP projection that acts as a complement of the preposition:

(9)


Assuming such a structure it is not at all clear how the NP or the PP should be moved stranding the relative behind. The PP that has to be moved contains both the NP and the relative clause. The same line of argumentation can be used for a case where the head-NP of the relative clause is a complement of another NP:

(10) *Ich habe [die Frau [DES Bürgermeisters]$_i$]$_{DP}$ gekannt, [der in unserer Stadt gelebt hat]$_i$.*
I have the wife the$_{GEN}$ mayor's known who in our town lived
'I have seen the mayor's wife, who has lived in our town

The head-NP (*des Bürgermeisters*) would move out of the relative clause adjoining to it and thereby forming an NP projection which acts as the complement of the dominating NP:

(11)


If the NP has to move to SpecAgrO (in order to receive case) and the relative clause is stranded behind it is unclear where the remaining relative should be attached to (if it were not by means of right-adjunction).

This is also true for the possessive construction found in the southern German dialects. Here the possessor NP is dative and presumably in a specifier position of the entire DP:

(12) *[$_{DP}$ [$_{DP}$dem Mann] [$_{D'}$ [$_{D°}$seine] [$_{NP}$Frau]]]*
the$_{DAT}$ man his wife

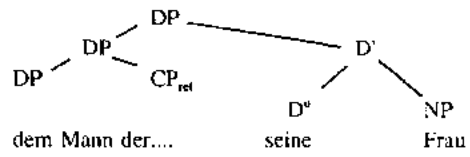
As has been mentioned already it is possible to construe a relative clause with the possessor phrase in SpecDP:

(13) (a) *Peter hat [[dem Mann den er nicht leiden konnte] seine Frau] geküßt.*
Peter has the$_{DAT}$ man, who he not stand could his wife kissed.
'Peter has the man who he couldn't stand's wife kissed

(b) *Peter hat [[DEM Mann]$_i$ seine Frau] [den er nicht leiden konnte]$_i$ geküßt.*
Peter has the$_{DAT}$ man his wife who he not stand could kissed

(c) *Peter hat [[DEM Mann]$_i$ seine Frau] geküßt [den er nicht leiden konnte.]$_i$*
Peter has the$_{DAT}$ man his wife kissed, who he not stand could

The possibility to extrapose the relative clause 'constituent-internally' (13b) or alternatively to a position following the verb creates a problem for the right-branching analysis. In this case we find the following base structure:

(14)


The DP (*der Mann*) moves out of the relative clause and adjoins to it. The relative clause should only be possible in a position preceding the possessive determiner. It is not obvious how it could be stranded behind. Another similar example is the following one:

(15) (a) *das Argument, daß pi irrational ist, das Peter präsentiert hat*
the argument that pi irrational is, which Peter presented has

(b) *das Argument, das Peter präsentiert hat, daß pi irrational ist*
the argument which Peter presented has, that pi irrational is

Here we find an NP that contains an argument clause. Presumably the entire NP is moved out of the relative clause. If one assumes that right-adjunction is not an option then one would have to assume for the sentence in (15) that the relative clause adjoins to CP. But in that case it is not clear anymore how the complement clause could be stranded behind as it is below:

(16) *Ich habe DAS Argument, das Peter präsentiert hat, gehört, daß pi irrational ist.*
I have the argument which Peter presented has, heard, that pi irrational is

Again the complement CP has no possible detachment site available if right-adjunction is excluded. All the examples discussed in the previous section constitute no problem for an analysis which makes use of right-adjunction, they all behave similar. (cf. section VI.2.3.)

2.3. The Nesting Requirement

Another difficulty for the stranding analysis is the constraint on the relative ordering of relative clauses discussed in section VI.2.2. (i.e. the nesting effect). Although this constraint does not have anything to do with movement it still creates similar problems concerning the predictions about constituenthood. If the subject precedes the object in the matrix clause the extraposed subject relative has to follow the extraposed object relative clause:

(17) (a) *daß jeder Mann eine Frau kennt, die rote Haare hat, der weiß wo's lang geht*
that every man a woman knows, who red hair has, who knows where to go
(b) **daß jeder Mann eine Frau kennt, der weiß wo's lang geht, die rote Haare hat*

A rightbranching analysis would predict just the wrong linear output (even if we assume an analysis which involves movement of the infinite verb to a higher position than the subject relative) as can be observed on basis of the following representation:

(18) $[_{CP}$ C° $[_{AgrSP}$ Subj $[_{AgrOP}$ Obj V° $[_{VP}$ RelSubj t_V $]_{V'}$ RelObj

2.4. No further stranding

The relative clause does not have to occur in the position following the verb but can move along with the NP. As noticed in Haider (1993) it is not at all clear why relative clauses cannot be stranded in the position the NP has to move to in order to receive case. According to Kayne (1993) stranding the relative clause should be possible in both positions. Data involving floated quantifiers show that in general it is not impossible to strand certain elements in this position (SpecAgrO):

(19) (a) *weil Peter viele Frauen geküßt hat.*
since Peter many women kissed has
(b) *Viele Frauen hat Peter geküßt.*
Many women has Peter kissed
(c) *[Frauen]$_i$ hat Peter [viele t$_i$] geküßt*
Women has Peter many kissed

The paradigm in (19) shows that a quantified NP can either be topicalized on its own (stranding the quantifier behind (19c) or the whole QP is topicalized (19b). There cannot be any principled reason why stranding in SpecAgrO should be prohibited. This can also be exemplified through the following paradigm:

(20) (a) *[Frauen]$_i$ hat Peter [viele t$_i$]$_j$ geküßt, [die er liebte]$_{ij}$*
women has Peter many kissed, who he loved
(b) *[Viele Frauen]$_i$ hat Peter geküßt [die er liebte]$_i$*
many woman has Peter kissed, who he loved
(c) *[Frauen]$_i$ hat Peter [viele t$_i$]$_j$ [die er liebte]$_{ij}$ geküßt*
women has Peter many who he loved kissed
(d) *[[Frauen]$_i$ [die er liebte]$_i$] hat Peter [viele t$_i$] geküßt*
women who he loved has Peter many kisse
(e) **[Viele [Frauen]$_i$]$_j$ hat Peter [die er liebte]$_{ij}$ geküßt*
many women has Peter who he loved kissed
(f) **[Frauen]$_i$ hat Peter t$_i$ [die er liebte]$_i$ geküßt*
women has Peter, who he loved kissed

Here we find a quantified NP modified by a relative clause. The relative clause can occur in extraposed position, no matter whether the quantifier is stranded (20a) or topicalized along with the NP (20b). It can also occur in a position adjacent to the stranded QP (20c) or adjacent to the topicalized NP (20d). However, the relative cannot occur in stranded position on its own (20e/f). (Notice that these data are also independent support for the present analysis. First, it shows that the relative clause

can either be associated with NP (X) or DP (x). This also means that we can interpret Split topicalization as another instance of making X available for identification. Second it also shows that the S-structure position of the identifyee determines the adjunction site for the identifier.

The same phenomenon can also be seen w.r.t. split topicalization.[1] Here we find an NP which is moved to SpecCP stranding some material (e.g. the determiner and an adjective) behind. Again the same paradigm concerning the possibility for stranding the relative clause can be observed:

(21) (a) *[Frau]$_i$ hat Peter [eine schöne t$_i$]$_j$ geküßt, [die blaue Augen hat]$_{ij}$.*
Woman has Peter a beautiful kissed, who blue eyes has

(b) *[Eine schöne Frau]$_i$ hat Peter geküßt [die blaue Augen hat]$_i$.*
A beautiful woman has Peter kissed, who blue eyes had.

(c) *[Frau]$_i$ hat Peter [eine schöne t$_i$]$_j$ [die blaue Augen hat]$_j$ geküßt*
Woman has Peter a beautiful, who blue eyes has kissed

(d) *[Frau]$_i$ [die blaue Augen hat]$_j$, hat Peter [eine schöne t$_i$]$_j$ geküßt*
Woman who blue eyes has, has Peter a beautiful kissed

(e) **[Frau]$_i$ hat Peter t$_i$ [die blaue Augen]$_i$ hat geküßt*
Woman has Peter, who blue eyes had kissed

Under a right-adjunction analysis these data follow immediately: S-structure serves as the input for the LCI. In case of Split Topicalization two possible identifyees are available. We can explain the ambiguity of the relative clause in extraposed position (whether it is interpreted as modifying the entire DP or just the (topicalized) NP. Both elements can serve as adjunction site for the relative clause. That fat that the relative clause cannot occur in stranded position is expected since in that case it is not adjoined to a projection dominating an identifyee. However, under the stranding analysis something needs to be said to exclude stranding in the specifier position of AgrOP.

2.5. An Ambiguity

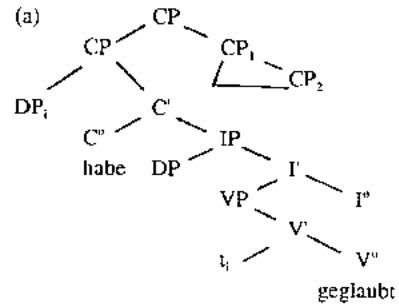
There is another empirical problem for the right-branching analysis concerning the predicted constituency the right-branching representation arrives at. Consider the following sentence:

[1] Here I will not go into a discussion of how split Topicalization is best analyzed (i.e. by means of movement or base-generation).

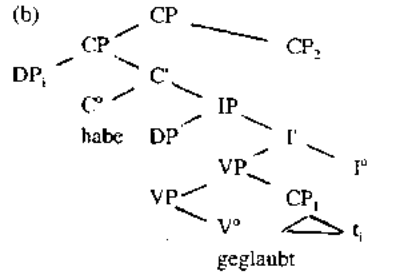
(22) *Die Tatsache$_{ji}$/DIE Tatsache$_{ji}$ habe ich geglaubt, [$_{CP_1}$daß Hans bedauert hat [$_{CP_2}$daß Maria gekommen ist.]$_i$]$_j$*
the fact/THE fact have I believed that John regretted has that Mary come is

There are two different readings available in case the determiner is contrastively stressed. The different readings correspond to two different representations. Under one reading (where no stress is necessary) the DP *die Tatsache* is identified by CP_1 (*daß Hans bedauert hat daß Maria gekommen ist*). The second reading corresponds to Identification of this DP only by the more deeply embedded CP_2 (*daß Maria gekommen ist*). According to our analysis this corresponds to the following two representations respectively

(23) (a)



(b)

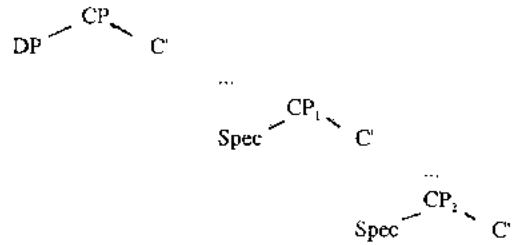


In (23a) the DP is identified by the entire CP_1 (including CP_2). The DP *the fact* is extracted out of the matrix clause. Therefore the interpretation arises that I believed the fact and the fact is *that John regretted that Mary has come*. In (23b) the DP is only identified by CP_2. In this case it is topicalized (successive cyclically) out of CP_2. It is interpreted as a complement of the verb *bedauern* rather than the matrix

verb *believe*. This results in the interpretation that I believed that John regretted the fact. The fact is *that Mary has come*.

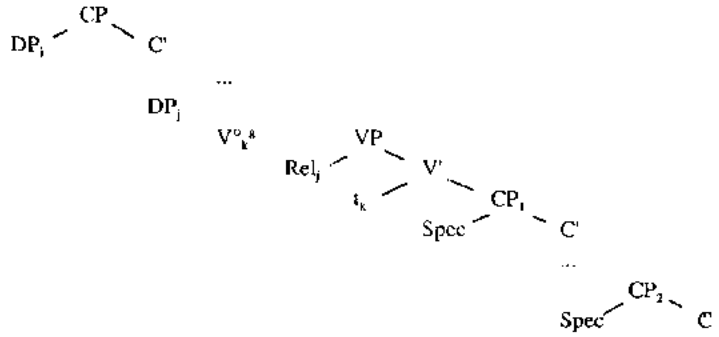
Now consider the representation an advocate of the right-branching analysis has to assume. It must be strictly right-branching in both cases. This means that the sentence receives the same representation for both readings.

(24)



The two different readings are a result of the two possible base-positions of the topicalized DP. As above it is either moved from a position within the matrix clause or it is moved out of the complement position of the verb within CP_1 (*bedauern*). On the basis of this example both analyses can derive the ambiguity. However, the difference is that under a right-adjunction analysis the two different readings correspond to two different phrase structures whereas in the right-branching analysis they are the same. This should make a difference. Suppose we add an extraposed relative clause which is correlated to a DP within the matrix clause. Now we expect the two analyses to make different predictions.

(25)

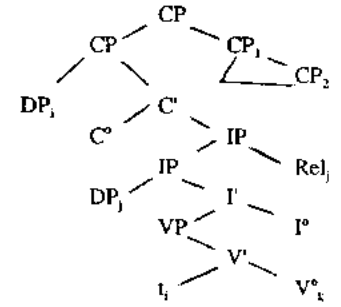


serialization: $DP_i\ DP_j\ V_k\ Rel_j\ CP_1\ CP_2$

[8] It has to be assumed that the verb is moved out of the VP.

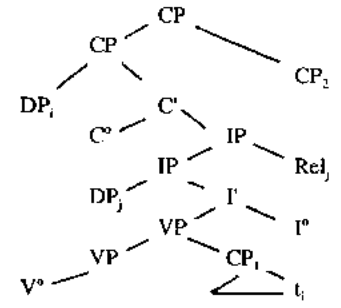
Under the right-branching analysis (which has to assume a stranding analysis for extraposed relative clauses) the representation would correspond to the one in (25). We expect the serialization indicated above (no matter which reading the sentence has, i.e. which CP is supposed to identify the topicalized DP). On the other hand the right-adjunction analysis makes different predictions. Here the position of the relative w.r.t. to the other constituents should differ with the different readings.

(26)



serialization: $DP_i\ DP_j\ V_k\ Rel_j\ CP_1\ CP_2$

(27)



serialization: $DP_i\ DP_j\ V_k\ CP_1\ Rel_j\ CP_2$

According to the right-adjunction analysis the serialization should differ w.r.t. to the two different readings whereas the right-branching analysis arrives at the same serialization for both readings. The serialization in (25), i.e. the only possible serialization for the right-branching analysis corresponds to the serialization in (26). According to the right-adjunction analysis this serialization necessarily results in the reading where the entire CP_1 (containing CP_2) identifies the DP. The second serialization (corresponding to Identification of DP by CP_2) is predicted to be ill-

formed under the right-branching analysis. However, it is predicted to be wellformed under the right-adjunction analysis.

Although the sentences are rather delicate to judge (as usual in case of multiple extraposition) the relevant example favors a right-adjunction analysis:

(28) (a) *Die Tatsache, hat der Mann$_j$ geglaubt, [der$_j$ Klara geküßt hat]$_{Rel_j}$ [$_{CP_1}$daß Peter bedauert hat] [$_{CP_2}$daß Maria gekommen ist]*
the fact has the man believed who Claire kissed has that Peter regretted has that Mary come is

(b) *DIE Tatsache$_i$ hat der Mann$_j$ geglaubt, [$_{CP_1}$daß Peter bedauert hat, [der$_j$ Klara geküßt hat]$_{Rel_j}$ [$_{CP_2}$daß Maria gekommen ist.]*
the fact has the man believed that Peter regretted has who Claire kissed has that Mary come is

Both predictions of the right-adjunction analysis are fulfilled. First, both serializations are well-formed. Secondly, the two different readings w.r.t. the DP *die Tatsache* correspond to the two different serializations - as expected under the right-adjunction analysis: In (28b) *DIE Tatsache* can just be interpreted as being identified by CP_2 (*daß Maria gekommen ist*). In (28a) however it is necessarily identified by the entire CP_1 *daß Peter bedauert hat, daß Maria gekommen ist.*

2.6. Coordination

Another set favoring the right-adjunction analysis involves VP-coordination. Consider the following sentence:

(29) *Er hat [vielen]$_i$ [[Gold anvertraut]$_{VP}$ oder [Geld ausgehändigt [denen ich nicht traue]$_i$]$_{VP}$]$_{VP}$*
He has [many]$_i$ [gold given]$_{VP}$ or [money given [who I not trust]$_i$]$_{VP}$
'He has to many people gold given or money given who I do not trust.'
(Haider, 1993)

It has to be noticed that according to Haider (1993) this sentence is ungrammatical. I do not share this judgement: to my informants this sentence does not even receive a question mark. Under a right-adjunction analysis this sentence is predicted to be grammatical since the head-NP determines the adjunction site of the relative clause. This means that in (29) the relative clause is adjoined to the first maximal projection dominating the identifyce. This means that it is also higher than the coordinated VP.

So the sentence in (29) does not cause a problem for the right-adjunction analysis (at least given the assumption that the sentence ins well-formed). Haider (who judges this sentence as ill-formed) argues that the ungrammaticality is predicted under the rightbranching analysis. It can be analyzed as an across-the-board violation: the second conjunct but not the first one contains a relative clause that is related to an antecedent that is not contained in either of the two conjuncts. However, this argumentation is only valid for Haider's analysis of extraposed relative clauses. He does not assume a stranding analysis for relative clauses. Therefore in the sentence above there is a relation between an element not contained within the conjuncts and the relative clause that is only contained in the second conjunct. This results in an across-the-board violation. However, it is not clear whether under a stranding approach towards extraposition (á la Kayne) the same argumentation is valid. Under such a view it could be argued that the antecedent *vielen* has been moved out of either conjuncts. In the first conjunct it would move out of its argument-position. In the second conjunct it would first move out of the relative clause and then from the argument-position further out of the conjunct. In this case the antecedent would have a relation to both conjuncts. Under this analysis the sentence in (29) might be possible to derive.

However, both under Haider's and Kayne's approach the following sentence should be perfectly well-formed (if the line of reasoning discussed above is adopted).

(30) **Er hat [vielen]$_i$ [[Gold ausgehändigt [denen er traute]$_i$]$_{VP}$ und [Geld gegeben [denen er mißtraute]$_i$]$_{VP}$]$_{VP}$.*
He has many gold given who he trusted and money given who he not trusted.

This prediction is not borne out: the sentence is ungrammatical. In (30) we are dealing with a coordinated VP. Here both conjuncts contain a relative clause (that is stranded behind in the complement position of the verb under the right-branching analysis). Both relative clauses are related to the head NP *vielen*, which is not contained in either of the conjuncts. This sentence does not constitute an across-the-board violation. As far as I can see under a right-branching analysis there is no obvious way to exclude this sentence. Under the right-adjunction analysis presented here this sentence is predicted to be ungrammatical. The antecedent that determines the adjunction site (at S-structure) is not contained in the VP. The relative clause must be adjoined to a position higher than the VP. Consider now the following paradigm:

(31) (a) *Maria hat [$_{VP}$[$_{VP}$[den Mann]$_i$ getroffen [den Peter kannte]$_i$]$_{VP}$ und [ihn geküßt]$_{VP}$]$_{VP}$*
Mary has the man met who Peter knew and him kissed

(b) *Maria hat [$_{VP}$[den Mann]$_i$ getroffen und ihn geküßt]$_{VP}$ [den Peter kannte]$_i$*
Mary has the man met and him kissed who Peter knew

The example in (31a) is predicted to be good under either analysis since the antecedent of the relative clause is contained within the conjunct. Under the right-branching analysis the relative clause is in complement position of the verb. The head NP moves to a position where it can receive case (we have to assume coordination of a projection higher than the VP in this case, but since this does not change the line of argumentation I will continue to use the term VP-coordination). Under the right-adjunction analysis the relative clause is adjoined to VP. It has to be assumed that adjunction of the relative clause does not change the status of the VP for coordination (i.e. a VP node created through adjunction can be coordinated with a simple one- segment VP). Under the right-adjunction analysis the sentence in (31b) is also expected to be grammatical since the relative can either be adjoined to the lower VP or to the VP-node created by coordination. However, the sentence in (31b) is expected to be ungrammatical under the right-branching analysis, nevertheless it is perfectly good. According to the right-branching analysis we would have to analyze (31b) as follows. We are dealing with a relative clause that follows the second conjunct. Therefore it should occupy the complement-position of the verb within the second conjunct. However, then the head of the relative clause would be a pronoun. For independent reasons pronouns can never act as the head NP of a restrictive relative clause, as exemplified below:

(32) **Maria hat [ihn]$_i$ geküßt [den Peter kannte]$_i$*
Mary has him kissed who Peter knew

To solve this problem within the right-branching analysis one could argue along the following lines: The head NP '*den Mann*' could move out of both conjuncts in an across-the-board manner. In that case the pronoun in the second conjunct could be analyzed as a resumptive pronoun. In this case we would deal with a structure that is similar to the one in (29). There the indirect object has been moved out of both of the conjuncts leaving a relative clause behind just in the second conjunct.[9]

[9] Notice that under such an analysis the right-adjunction approach maintained here could still explain the grammaticality of (31b) since then the antecedent (above the coordinated constituent) is expected to determine the adjunction site).

However, there are some problems with the assumption that we are dealing with a resumptive pronoun. First it needs to be explained that the first conjunct does not allow for a resumptive pronoun:

(33) *Maria hat den Mann [$_{VP}$[$_{VP}$(*ihn) getroffen]$_i$ und [$_{VP}$(ihn) geküßt]]$_{VP}$.*
Mary has the man (*him) met and (him) kissed

Secondly, consider the following contrast:

(34) (a) *?Den Mann$_i$ hat Maria [t$_i$ getroffen]$_{VP}$ und [ihn geküßt]$_i$ [den Peter kannte]$_i$.*
The man has Mary met and him kissed who Peter knew

(b) **Den Mann$_i$ hat Maria [getroffen [den Peter kannte]$_i$]$_{VP}$ und [ihn geküßt]$_{VP}$.*
The man has Mary met who Peter knew and him kissed

In these examples the head of the relative clause is moved out of the first conjunct whereas a (coreferent) pronoun is still within the second conjunct. If one does not assume that the pronoun is a resumptive pronoun both sentences are instances of an across the board-violation.[10] For the moment let me abstract away from this problem. It has to be noticed that there is a contrast in (34). This contrast is only expected under the right-adjunction analysis. There is no possible adjunction site for the relative clause in (34b) since the identifyee occupies SpecCP. Therefore the relative clause has to be adjoined to CP as well. There is no obvious reason for the right-branching analysis to exclude (34b) and at the same time allow for (34a). Besides the across-the-board violation there is no difference between (34a) and (34b).

2.7. Deletion

VP-Deletion constitutes one of the traditional tests for where an extraposed element is adjoined to. We have already discussed the problems concerning the analysis. However, it is important to notice that sentences like (35-36) are not quite straightforward to account for in a right-branching analysis (no matter what sort of analysis one assumes for VP-deletion).

[10] The fact that (34a) is not perfectly well-formed can either be due to an across-the-board violation or to the fact that German does not allow for resumptive pronouns.

(35) *John met a man last week (who was) from Philadelphia, and George (met a man last week/*did) (who was) from New York.*

(36) *A MAN came in with BLOND hair, and a WOMAN did (e) with BROWN hair.*

Under a rightbranching analysis the ungrammaticality of VP-ellipsis with an object relative (which is itself not deleted) is expected. On the other hand it is not expected that a subject relative can still be overtly present. If the linearly last element is the most deeply embedded element then deletion of the VP (containing it) should obligatorily delete the relative clause as well. On the other hand the right-adjunction analysis can equally account for both sentences. The adjunction site is determined by the identifyees. The relevant identifiers have to be adjoined to a position higher than the deletion site.

3. Specificity and the Mapping Hypothesis

There is another problem with the assumption that the VP in German is head-initial. Notice the issue I am discussing now is highly problematic since there is no consensus about the analysis one has to assume for the facts I am interested here. Still, I think that it is worth to discuss it briefly. The problem concerns the Mapping Hypothesis (á la Diesing 1992). It is claimed there that the position of NPs maps to different interpretations. The assumption is that non-specific NPs stay within VP whereas specific NPs move out of the VP in the specifier position of a functional projection.[11] The difference is exemplified by the following sentences:

(37) (a) *Peter hat der Maria ja doch ein Buch gegeben.*
Peter has the Mary prt. a book given

(b) *Peter hat der Maria ein Buch ja doch gegeben.*
Peter has the Mary a book prt. given

The particle *ja doch* is assumed to mark the VP boundary. Therefore in (37a) the NP which is interpreted as non-specific has to be within the VP. In (37b) where *ein Buch* is interpreted as specific the Adverb follows it. So one can conclude that it is moved out of the VP.

[11] Cf. e.g. Haiden (1995) for a criticism of Diesing's version of the Mapping Hypothesis.

Of course if the assumption that the interpretation of NPs corresponds to their syntactic position it cannot be true that the VP in German is head-initial. If both claims would be true under their simples version we would expect non-specific NPs to occur post-verbally:

(38) **Peter hat der Maria gegeben ein Buch*
Peter has the Mary given a book

This is however not a possible option. It seems that we have to give up either the Mapping hypothesis (in the version presented by Diesing 1992) or the assumption that the German VP is head-initial. Zwart (1992) takes up the first line of reasoning (giving up the simple Mapping Hypothesis). Theoretically he argues that syntax and semantics should be kept strictly separated, i.e. NPs are moved to SpecAgrO in order to receive case. There is no semantic trigger to block movement. His empirical argument is that there is no adjacency requirement between a non-specific NP and the verb:

(39) (a) *dat Jan vaak meisjes kust*
that Jan often girls kisses

(b) *dat Jan meisjes vaak kust*
that Jan girls often kisses

Several other elements (analyzed as Small Clause predicates) have to be adjacent to the verb:

(40) (a) *dat Jan Marie op (*gisteren) belde* (particles)
that Jan Mary up yesterday called

(b) *dat Jan de deur rood (*gisteren) verfde* (resultatives)
that Jan the door red yesterday coloured

(c) *dat Jan de sloot in (*gisteren) sprong* (directional PP)
that Jan the chimney in yesterday jumped

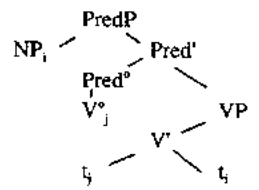
(d) *dat het lijk in de kast (*gisteren) zat* (locational PP)
that the body in the cupboard yesterday sat

On the other hand Zwart (1992) argues independently that the verb undergoes short movement in Dutch. This can be shown since adjuncts can occur between the verb and a complement clause:

(41) *Wat wil je dat ik zeg op de vergadering dat ik von je voorstel vind*
What want you that I say at the meeting dat I of your proposal think

If the verb can be shown move in one Dutch construction then, by minimalist assumptions, it has to move in any other Dutch sentence, too. Therefore Zwart concludes that the adjacency requirement for Small Clause predicates is a result of a Spec Head relation:

(42)



Notice that for this line of argumentation it is crucial to assume that movement is never optional. The obligatory movement of the Small Clause Predicate is responsible for the fact that they can never occur in a position following the verb (a fact for which Zwart argues that it would have to be stipulated within a right-adjunction analysis).

(43) (a) **dat Jan Marie belt op*
that Jan Mary called up
(b) **dat Jan de deur verft rood*
that Jan the door coloured red
(c) **dat Jan springt de sloot in*
that Jan jumps the chimney in
(d) **dat het lijk zit in de kast*
that the body sits in the cupboard

To summarize: on the one hand the interpretation of specificity is rejected as a reflex of syntactic structure. On the other hand the observed SOV order is assumed to be derived via short movement of the verb to a position immediately dominating VP. Certain NPs (in particular Small Clause Predicates) have to move into SpecPredP resulting in a strict adjacency requirement between them and the verb. Other NPs invariably move to SpecAgrO in order to receive case. A consequence of this approach is that one has to assume that adverbs can be generated in several positions. Therefore the position of adverbs w.r.t. NPs cannot be taken as a test for the actual position of the relevant NPs.

In the sentences below (repeated here for convenience) rather than to assume that the adverb shows whether the NP has moved it is assumed that it is the adverb which is adjoined to two different positions. The actual position of the adverb is

interpreted to be responsible for the difference in interpretation.

(44) (a) *dat Jan vaak meisjes kust*
that Jan often girls kisses
(b) *dat Jan meisjes vaak kust*
that Jan girls often kisses

Koster (1993) argues for the head-initial hypothesis without giving up the idea that specificity is a reflex of the syntactic position. He argues that so called *weak NPs* have to move to PredP. This is assumed to be some sort of *incorporation*. Due to this incorporation the nonspecific reading occurs. In a way he assumes a different concept of specificity than Diesing does. Under his approach non-specificity corresponds to a *property* reading. Non-specific NPs do not have argument status. Therefore they have to incorporate. Specific NPs on the other hand have real argument status and are therefore interpreted as being referential. This referentiality induces the specific reading. They can move to an argument position, namely SpecAgr.

The distinction between the specific and the non-specific interpretation in terms of referential vs. non-referential is reminiscent of Stowell's (1987) idea that NPs do not refer whereas DPs do. Therefore the incorporation hypothesis is motivated also semantically since a property cannot saturate the argument position of the verb. It has to combine with the verb to form a complex predicate.

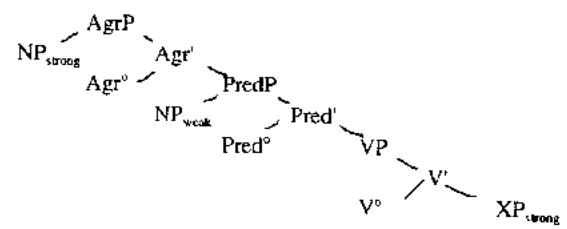
Let us compares Diesing's claim that non-specific NPs stay in VP and Koster's claim that they move to SpecPredP via incorporation. Here difference between the two approaches disappears. Staying within the VP can also be interpreted as some particular instance of incorporation.[12] NPs that denote a property cannot move out of the VP. They can be assumed to form a complex VP. The necessity to assume another functional projection (PredP) seems to be a result of the assumption that the VP is head-initial. In order to derive the adequate word-order for Dutch (and German) one has to assume that the NP still moves out of the VP (although incorporation to the verb is still maintained).

The existence of extraposition induces the following problem for Koster's approach: He assumes that weak NPs have to move to SpecPredP whereas other NPs have to move to AgrP. Weak NPs therefore constitute a semantic unit with the

[12] Cf. Borer (1994) as well as Haiden (1995) for a similar approach.

verb. Strong NPs do not have this property. The problem is that elements that appear in a position following the verb (i.e. extraposed constituents) are assumed to stay in situ (i.e. in the complement position of the verb). According to the above line of reasoning it would be a natural assumption that these elements also receive a weak reading. However CPs in extraposed position can be interpreted as referential (saturated) arguments. The following paradigm occurs:

(45)



Weak NPs in SpecPredP constitute a semantic unit with the verb. It therefore seems strange that elements which stay in the VP can be strong. It would be more natural to expect them to have the same interpretation as the weak ones.

However, under an analysis where it is assumed that elements following the verb are in a position adjoined to VP it is straightforward that they pattern with elements that occur in a position above the VP. All that has to be said is that weak NPs stay in VP (probably being incorporated) and strong XPs are not contained in the VP (no matter whether they are moved to a Specifier position to the left of the VP like it is the case for NPs or whether they are adjoined to a position above the VP as it is the case for CPs and PPs). As far as I can see the latter approach has two advantages: First, there is no need to postulate a projection that is only there to capture the intuition that there is a kind of incorporation process. Assuming that elements staying in the VP receive this interpretation is a more natural assumption. Second the fact that elements in 'extraposed' position pattern with elements that are not incorporated follows more easily since they are not within the VP. The right-branching analysis has to stipulate this pattern.

4. Antisymmetry, Right-adjunction and Precedence

Let me now briefly discuss my proposal in the light of Kayne's (1993) antisymmetry proposal. I will concentrate on two main points[13]:

- The main claim in Kayne (1993) is that syntactic structure unambiguously determines linear ordering. Asymmetric c-command maps into linear precedence. From this assumption, it follows that right-adjunction is excluded.
- This proposal crucially relies on the assumption that any ordering relation has to be asymmetric to be a total ordering.

As we have seen there is strong empirical support for the assumption that extraposition is right-adjunction. I have tried to show that this is also theoretically supported by the theory of licensing and the interaction of linear ordering and hierarchical structure (cf. chapter I.)

If one compares the proposal I have made in this dissertation and the two main assumptions of Kayne's proposal, there are two points to be highlighted. First, I have argued that there is a precedence constraint for Identification. Secondly, Identification was shown to be an asymmetric relation (in fact, it involves two asymmetric relations, one w.r.t. inherent licensing, and another one w.r.t. relational licensing, i.e. for determining the content of a given element). I would like to suggest that one can still maintain the claim that syntactic structure unambiguously determines linear ordering. However, asymmetric c-command is not the only device. In case of Identification there is an independently motivated device for determining precedence.[14]

Furthermore, the locality constraint assumed for Identification receives an interesting interpretation under such a view. First, notice that for the definition of i-government I follow Kayne (1993) in assuming that c-command is only visible for categories. However, I do not assume that dominance is defined in terms of categories. I propose the following definition of c-command:

[13] I will not discuss Kayne's proposal in any detail.

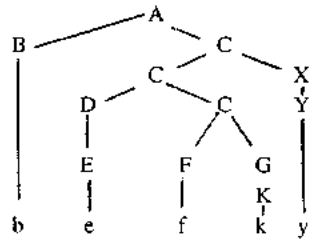
[14] Notice that this is more or less in the spirit of the proposal by Brody (1994), who assumes that dependency relations in general map into precedence as well as Williams' (1994) proposal that dependency is governed by precedence whereas coreference is governed by c-command.

(46) c-command:
X c-commands Y iff X & Y are categories and no instance of X dominates Y and every node that dominates X dominates Y.

This definition differs from Kayne's in that it allows for multiple adjunction. It seems to me that this assumption is theoretically superior[15] because of the following reason: First, if one adds the condition that c-command holds only for categories, the statement that dominance holds only for categories becomes somehow redundant because c-command is defined in terms of dominance.

Technically, I can integrate Kayne's proposal into mine in the following way. In the tree in (47) X corresponds to the right-adjoined identifier. D corresponds to the identifyee that is specified for precedence:

(47)



According to Kayne, asymmetric c-command can be mapped into a set of ordered pairs. The first element of each of these pairs asymmetrically c-commands the second one, i.e. the first member has to linearly precede the second). The structure in (47) derives the following ordered pairs:

(48) <B,X>, <B,Y>, <B,D>, <B,E>, <B,F>, <B,G>, <B,K>, <u><X,D>, <X,E>, <X,F>, <X,G>, <X,K></u> <D,F>, <D,G>, <D,K>, <F,K>

The underlined pairs constitute the problematic pairs. In all these pairs the first member asymmetrically c-commands the second one. However, the linear output we derive shows that the first member follows the second member. Therefore Kayne has to exclude this configuration. Here I would tentatively suggest the following proposal. One could assume that if X asymmetrically c-commands D, but D is specified for precedence, then all the ordered pairs containing the element that

[15] It seems to me that it also empirically superior; but it is not the purpose of this paper to discuss this in more detail.

induces this contradiction are reverted. This is stated below in (49). If true, then (50) is the set of ordered pairs that corresponds to the structure in (47) resulting in the set of ordered terminal pairs (51):

(49) If A asymmetrically c-commands B (resulting in the ordered pair <A,B>) but the PCI requires that B has to precede A then all the ordered pairs involving B are reversed.

(50) <B,X>, <B,Y>, <B,D>, <B,E>, <B,F>, <B,G>, <B,K>, <u><D,X>, <E,X>, <F,X>, <G,X>, <K,X></u> <D,F>, <D,G>, <D,K>, <F,K>

(51) d(A): <b,y>, <b,e>, <b,f>, <b,k>, <e,y>, <e,f>, <e,k>, <f,k>, <f,y>, <k,y>

Assuming (49) one can also derive that the problematic (right-adjoined) element has to be higher than the element that is specified for precedence. Otherwise, the contradiction triggering the reversal could never occur. If the identifier would be lower than the element specified for precedence then the latter would asymmetrically c-command the identifier and no contradiction would occur that would trigger (49) to apply.

On the other hand, the locality domain can be defined as follows: the element that is specified for precedence has to be the *first element to be c-commanded*. Note that this is a definable relation. Every c-command domain (a domain we can also define as a set) is itself partially ordered by the dominance relation. In a given c-command domain of a given element *the first element to be c-commanded* is the element that is dominated by the least nodes possible.

I will not pursue this issue any further here. The last section was simply intended to give a brief outline of an analysis that can combine the possibility for right-adjunction and Kayne's main proposal at the same time. A further elaboration of such a view, and the exact implementation is not the goal of the present analysis.

CONCLUSION AND PROSPECTS

In this dissertation I tried to develop an empirically and theoretically adequate analysis for extraposition. For doing so, I introduced the notion of *Identification at Domain D*. This notion makes it possible to derive several properties of extraposition, especially the obligatory rightness and the distinction between left- and rightward relations are derived without referring to 'rightness'. Furthermore the analysis developed here makes it possible to implement various different presuppositions associated with different kinds of Identification. This allows us to take discourse into account without giving up the Autonomy-of-Syntax-Hypothesis.

Several apparent problems for the present analysis forced us to look into areas that are not immediately related to the core cases the analysis is intended to capture. I have tried to develop an analysis for the syntax and semantics of NPs, optionality, conditionals, result clauses and Left Dislocation etc. In many cases I could only sketch the direction into which I think that the present analysis leads us to analyze various of these phenomena. However, I did not address all questions that the analysis induces. I think that the notion of Identification is relevant in other contexts as well, i.e. it is not only a relation that is intended to capture the properties of extraposition. I argued that Identification is a primitive (licensing) relation within the Theory of Grammar, and I have argued that it is not only relevant at domain D but also at LF (binding) and presumably S-structure (ECP, licensing of empty categories). This means that it can be instantiated in various different ways. There are several other areas where I think that the notion of Identification (at domain D) can lead to some interesting conclusions.

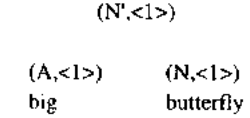
1) Copula sentences.
Copula sentences seem to be an instance of Identification. It is instantiated differently and therefore we expect different constraints. Identification is overtly realized by means of the copula. Let me just mention an interesting parallelism to the instances I have analyzed in this dissertation. There is a crucial property of (identificational) copula sentences (w.r.t. domain D) that is quite similar to the cases of Identification discussed here. It is not true that 'identificational' copula sentences involve a symmetric relation. Therefore, we cannot simply revert the two NPs without any difference in meaning. This can be seen in the following question-answer pairs:

(1) Q: *Wer ist Dr. Watson?*
Who is Dr.Watson?
A1: *[Dr. Watson] ist [der Gehilfe von Sh. Holmes]*
'Dr.Watson is Sh.Holmes' assistant'
A2: *#[Der Gehilfe von Sh.Holmes] ist [Dr. Watson]*
'Sh.Holmes' assistant is Dr.Watson'

(2) Q: *Wer ist der Gehilfe von Sh.Holmes?*
'Who is the assistant of Sh.Holmes?'
A1: *#[Dr. Watson] ist [der Gehilfe von Sh. Holmes]*
A2: *[Der Gehilfe von Sh.Holmes] ist [Dr. Watson]*

With neutral intonation there is only one appropriate answer for the questions in (1) and (2) respectively. The generalization we find is predicted by the Precedence Constraint of Identification: The identifyee precedes the identifier.

2) Theta-Identification and Modification.
Another instance where I think that Identification is relevant is what Higginbotham (1985) calls "theta-identification" (it might turn out that this is an instance of Identification at D-structure). Notice that I have sketched the parallelism between theta-identification and our instances of Identification. Consider Higginbotham's analysis for a phrase like *big butterfly*:

(3)


The slots in the argument structures of A and N (<1>) are identified. There is an interesting parallelism and interesting predictable difference to our instances of Identification. The parallelism is the locality constraint: Higginbotham assumes that theta-identification takes place under government a very local domain - similar to our locality constraint. There are however various differences to our instances of Identification. Among those differences are the following ones

a) there is no coindexation between the identifyee and the identifier

This means that the Precedence Constraint cannot apply. The same is true for instances of (theta-) Identification between an overt and a covert position. (Here we have only sketched some properties w.r.t. extraposed clauses without an overt correlate in the matrix clause. The exact nature of this relation still has to be analyzed more carefully.

b) providing a restriction by an adjective does not suffice to satisfy the obligatory need of Identification:

(4) *DER große Mann,*(der gerne Bier trinkt)*
THE big man *(who likes beer drinking)

However, for introducing a discourse referent such cases of Identification by an adjective seems to suffice:

(5) Q: *Hat Peter ein Buch gelesen?*
Has Peter a book read?
A1: *#Ja, Peter hat das Buch gelesen?*
Yes, Peter has the book read
A2: *Ja, Peter hat das vorgestern erschienene Buch gelesen.*
Yes, Peter has the the-day-before-yesterday published book read.

One could assume that adjectives are always related to N (i.e. the property X but not to D (x). I think that many interesting issues arise in discussing the exact nature of those differences.

3) Cross-linguistic variation.
English seems to differ from German in several different respects. It seems to me that many of them are related. This means that we have to look for the parametric choices Identification makes possible. These can concern lexical properties of determiners and pronouns, the Locality Constraint etc.

Moreover, it will be interesting to look into strict head-final languages where the identifyee follows the identifier. It seems to me that this should not lead to the conclusion that the Precedence Constraint should be a possible parametric option. I think that apparent violations of the Precedence Constraint should be analyzed in a different way, depending on the properties of the two elements that instantiate the relation of Identification. These examples seem to suggest that it could rather be a matter of theta-Identification, i.e. parallel to adjectival modification.

4) Non-restrictive Modification.
As I briefly sketched in Appendix II. we expect non-restrictive modification to behave differently. It is not an instance of Identification. However, there are many resemblances which could lead to some interesting generalizations on the difference between restrictive and non-restrictive modifications. The difference seems to be

dependent on the fact that we do not find a subset-relation (as in the cases we have been discussing) but a pure 'identity-relation'. These are only some of the areas that occur to me as being worth looking at under the light of the present analysis.

References:

Abney, S. (1987) *The English noun phrase in its sentential aspects*. PhD diss., Cambridge, Mass.: MIT.

Akmajian, A. (1975) "More evidence for an NP-cycle." *Linguistic Inquiry* 6, 115-129.

Akmajian, A.; R. Jackendoff (1970) "Coreferentiality and Stress." *Linguistic Inquiry* 1, 124-126.

Altmann, H. (1981) *Formen der Herausstellung im Deutschen. Rechtsversetzung, Linksversetzung, Freies Thema und verwandte Konstruktionen*. Tübingen: Niemeyer.

Anagnostopoulou, E. (1994) *Clitic Dependencies in Modern Greek*. PhD diss., Salzburg.

Asakawa, T. (1979) "Where does the Extraposed Element move to?" *Linguistic Inquiry* 10, 505-508.

Asher, N. (1993) *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers.

Bach, E. (1974) *Syntactic Theory*. New York: Holt, Rinehart & Winston.

Baltin, M. (1975) "On the Cyclicity of Extraposition." *Linguistic Inquiry* 6, 465-468.

Baltin, M. (1978) *Toward a Theory of Movement Rules*. PhD diss. Cambridge, Mass.: MIT.

Baltin, M. (1981) "Strict Bounding." C.L. Baker; J. McCarthy (eds.) *The Logical Problem of Language Acquisition*. Cambridge, Mass.: MIT Press.

Baltin, M. (1982) "A landing site theory of movement rules." *Linguistic Inquiry* 13, 1-38.

Baltin, M. (1983) "Extraposition: Bounding vs. Government-Binding." *Linguistic Inquiry* 14, 155-162.

Baltin, M. (1984) "Extraposition Rules and Discontinuous Constituents." *Linguistic Inquiry* 15, 157-163.

Baltin, M. (1987) "Do Antecedent Contained Deletion Exist?" *Linguistic Inquiry* 18, 579-595.

Bayer, J. (1991) *Directionality of Government and Logical Form. A Study of Focusing Particles and Wh-Scope*. Habilitationsschrift, Universität Konstanz.

Beneš, E. (1968) "Die Ausklammerung im Deutschen als grammatische Norm und als stilistischer Effekt." (reprinted in:) P. Braun (ed.) *Deutsche Gegenwartssprache*. München: Wilhelm Fink Verlag, 321-338.

Bennis, H. (1986) *Gaps and Dummies*. Dordrecht: Foris.

Borer, H. (1994) Lecture given at the Girona Summer School in Linguistics (GISSL 1994).

Bouchard D. (1984) *On the Content of Empty Categories*. Dordrecht: Foris.

Brody, M. (1984) "On Contextual Definitions and the Role of Chains." *Linguistic Inquiry* 15, 355-380.

Brody, M. (1990) "Some remarks on the focus field in Hungarian" *UCL working papers in linguistics 2*, UCL.

Brody, M. (1994) "Phrase Structure and Dependence." *Working Papers in the Theory of Grammar*, Vol.1, No.1, Budapest: Research Institute for Linguistics, Hungarian Academy of Science.

Büring, D.; K. Hartmann (1994) "Doing the Right Thing - Extraposition as a Movement Rule." *Sprachwissenschaft in Frankfurt* Arbeitspapier Nr. 13.

Calabrese, A. (1990) "Some Remarks on Focus and Logical Structure in Italian", *Harvard Working Papers in Linguistics*.

Campos, H. (1986) "Indefinite Object Drop." *Linguistic Inquiry* 17, 354-359.

Cardinaletti, A. (1990) "Subject/object asymmetries in German null-topic constructions and the status of SpecCP." J. Mascaró, M. Nespor (eds.) *Grammar in Progress. Glow Essays for Henk van Riemsdijk*. Dordrecht: Foris, 75-84.

Cardinaletti, A. (1990) "Es, pro and Sentential Arguments in German." *Linguistische Berichte* 126, 135-164.

Chater, N.; M. Pickering; D. Milward (1994) "What is Incremental Interpretation?" ms. Center for Cognitive Science, University of Edinburgh.

Chierchia, G. (1992) "Anaphora and Dynamic Binding." *Linguistics and Philosophy* 15, 111-183.

Chomsky, N. (1975) *Reflections on Language*. New York: Pantheon.

Chomsky, N. (1976) "Conditions on rules of Grammar." *Linguistic Analysis* 1, 75-109.

Chomsky, N. (1977) "On Wh-movement." P. Culicover, T. Wasow and A. Akmajian (eds.) *Formal Syntax*. New York: Academic Press.

Chomsky, N. (1980) "On Binding." *Linguistic Inquiry* 11, 1-46.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986a) *Knowledge of Language*. New York: Praeger.

Chomsky, N. (1986b) *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1992) "A Minimalist Program for Linguistic Theory." *MIT Occasional papers in Linguistics*. 1, Cambridge, Mass., MIT.

Chomsky, N. (1994) "Bare Phrase Structure." *MIT Occasional Papers in Linguistics* 5, Cambridge Mass., MIT.

Chomsky, N. (1995) "Chapter 4 Draft." ms. Cambridge, Mass., MIT

Cinque, G. (1977) "The Movement Nature of Left Dislocation." *Linguistic Inquiry* 8, 397-412.

Cinque, G. (1990) *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.

Clark, R. (1990) "Scope Assignment and Modification." *Linguistic Inquiry* 23, 1-28.

Culicover P.; M. Rochemont (1990) "Extraposition and the Complement Principle." *Linguistic Inquiry* 21, 23-47.

Culicover, P.; M. Rochemont (1994) Paper given at the Tilburg Conference on Rightward Movement.

Cummings, R. (1984) "Functional Analysis." E. Sober (ed.) *Conceptual Issues in Evolutionary Biology*. Cambridge, Mass.: MIT Press, 386-407.

Dik, S.C. (1978) *Functional Grammar*. Amsterdam: North Holland Publishing Company.

Diesing, M. (1992) *Indefinites*. Cambridge, Mass.: MIT Press.

di Sciullo, A.-M.; E. Williams (1987) *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Donellan, K. (1966) "Reference and Definite Description." *The Philosophical Review* 75, 281-304.

Donellan, K. (1978) "Speaker Reference, Descriptions and Anaphora." P. Cole (ed.) *Syntax and Semantics Vol. 9: Pragmatics*. New York: Academic Press, 47-68.

Doron, E. (1983) *Verbless Predicates in Hebrew*. PhD diss., University of Texas at Austin.

Downing, B. (1978) "Some Universals of Relative Clause Structure." J.H. Greenberg (ed.) *Universals of Human Language Vol. 4, Syntax*. Stanford, California: Stanford University Press, 375-418.

Drubig, H. (1992) "On Topicalization and Inversion." *Arbeitsberichte des Sonderforschungsbereichs 340: Sprachtheoretische Grundlagen für die Computerlinguistik* 22.

Emonds, J. (1976) *A Transformational Approach to English Syntax. Root, Structure-Preserving and Local Transformations*. New York: Academic Press.

Enç, M. (1991) "The Semantics of Specificity" *Linguistic Inquiry* 22, 1-25.

Fauconnier, G. (1975) "Pragmatic Scales and Logical Structure" *Linguistic Inquiry* 6, 353-375.

Fiengo, R. (1986) "Definiteness, Specificity and Familiarity." *Linguistic Inquiry* 18, 163-166.

Fiengo, R.; R. May (1994) *Indices and Identity*. Cambridge, Mass.: MIT Press.

Giusti, G. (1991) *La sintassi dei sintagmi nominali quantificati. Uno Studio comparativo*. PhD diss., University of Venice & Padua.

Grimshaw, J. (1990) "Extended Projections", ms. Brandeis University.

Guerón, J. (1976) "Interpretation of PP complements: a semantic filter for PP extraposition." H.v. Riemsdijk (ed.) *Green ideas blown up*. Papers from the Amsterdam Colloqium of Trace Theory. Institut for Algemene Taalwetenschap, University of Amsterdam.

Guerón, J. (1980) "The Syntax and Semantics of PP-Extraposition." *Linguistic Inquiry* 11, 637-678.

Guerón, J.; R. May (1984) "Extraposition and Logical Form." *Linguistic Inquiry* 15, 1-31.

Haiden, M. (1994) "Split Topicalization." *Wiener Linguistische Gazette* 48-50, 31-60.

Haiden, M. (1995) "On the Semantic Evaluation of NP-syntax." MA-thesis. University of Vienna.

Haider, H. (1993) "Detached Clauses - The latter the deeper." ms. Univ. Stuttgart

Haider, H. (1995) "Extraposition." ms. Univ. Stuttgart

Haider, H. (1994) "Downright down to the right." ms. Univ. Stuttgart

Hawkins, J. (1978) *Definiteness and Indefiniteness*. London: Croom Helm.

Hegarty, M. (1991) *Adjunct Extraction and Chain Configurations*. PhD diss., Cambridge, Mass., MIT.

Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*. PhD. diss., Univ. of Mass., Amherst.

Ladusaw, W. (1983) "Logical Form and conditions on grammaticality." *Linguistics and Philosophy*, 6: 373-392.
Larson, R. (1988) "On the Double Object Construction." *Linguistic Inquiry* 19, 335-391.
Lasnik, H. (1991) "On the Necessity of Binding Conditions." R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge, Mass., MIT Press, 7-28.
Lasnik, H.; M. Saito (1984) "On the Nature of Proper Government." *Linguistic Inquiry* 15, 235-289.
Lasnik, H.; M. Saito (1992) *Move alpha*. Cambridge, Mass.: MIT Press.
Lasnik, H.; T. Stowell (1991) "Weakest Crossover." *Linguistic Inquiry* 22, 687-720.
Lakoff, G. (1968) "Pronouns and Reference." mimeograph, Indiana University Linguistics Club, Indiana University, Bloomington, Indiana.
Lehmann, Ch. (1979) *Der Relativsatz. Typologie seiner Strukturen. Theorie seiner Funktionen. Kompendium seiner Grammatik*. (= Arbeiten des Kölner Universalien - Projekts, 36)
Liberman, M. (1974) "On Conditioning the Rule of Subject-Auxiliary Inversion." E. Kaisse; J. Hankamer (eds.) *Papers from the Fifth Annual Meeting of the North Eastern Linguistics Society*. Cambridge, Mass.: Harvard University.
Lieber, R. (1983) "Argument Linking and Compounding in English." *Linguistic Inquiry* 14, 251-286.
Löbel, E. (1990) "D und Q als funktionale Kategorien in der Nominalphrase." *Linguistische Berichte* 127, 232-264.
Mac Cawley, J. (1993) *Everything that Linguists have Always wanted to Know about Logic (but were ashamed to ask)*. Chicago: The University of Chicago Press.
Marillier, J.F. (1993) "Integration vs. Extraposition von Relativsätzen." J.F. Marillier (ed.) *Satzanfang-Satzende. Syntaktische, semantische und pragmatische Untersuchungen zur Satzabgrenzung und Extraposition im Deutschen*. Tübingen: Gunter Narr, 219-232.
May, R. (1985) *Logical Form: Its Structure and Derivation*. Cambridge, Mass.: MIT Press.
Montague, R. (1970) "Universal Grammar." R. Thomason (ed.) (1974), 222-246.
Müller, G. (1994) *On Extraposition and Successive Cyclicity*. Ms. Univ. Tübingen
Müller, G.; W. Sternefeld (1993) "Improper Movement and Unambiguous Binding." *Linguistic Inquiry* 24, 461-507.
Nakajima, H. (1992) "Another Type of Antecedent Government." *Linguistic Inquiry* 23, 313-328.
Neale, S. (1990) *Descriptions*. Cambridge, Mass.: MIT Press.
Olsen S. (1991) "Die deutsche Nominalphrase als "Determinerphrase". In: S. Olsen; G. Fanselow (eds.) *DET, COMP, und INFL. Zur Syntax funktionaler Kategorien und grammatischer Funktionen*. Tübingen: Niemeyer: 35-56
Perlmutter, D.; J. Ross (1970) "Relative Clauses with Split Antecedents." *Linguistic Inquiry* 1, 350.
Pesetsky, D. (1982) *Paths and Categories*. PhD. diss., Cambridge, Mass. MIT.
Pesetsky, D. (1985) "Morphology and Logical Form." *Linguistic Inquiry* 16, 193-246.
Pesetsky, D. (1987) "Wh-in-situ: Movement and Unselective Binding." E.J. Reuland; A.G.B. ter Meulen (eds.) (1987), 98-129.
Pesetsky, D. (1990) "Experiencer Predicates and Universal Alignment Principles," ms. Cambridge, Mass., MIT.
Pesetsky, D. (1992) "Zero Syntax". ms. Cambridge, Mass., MIT
Postal, P.; G. Pullum (1988) "Expletive Noun Phrases in Subcategorized Positions." *Linguistic Inquiry* 19, 635-670.
Prince, E. (1991) "Syntax and Discourse: A Look at Resumptive Pronouns," ms. University of Pennsylvania.
Prince, E. (1992) "The ZPG letter: subjects, definiteness, and information-status." S. Thompson; S. and W. Mann (eds.) *Discourse description: diverse analyses of a fund raising text*. Philadelphia/Amsterdam: John Benjamins B.V., 295-325.
Prinzhorn, M. (1992) *Funktionale Projektionen*. PhD diss., University of Vienna.
Prüst, H.; R. Scha; M. van den Berg (1994) "Discourse Grammar and Verb Phrase Anaphora." *Linguistics and Philosophy* 17, 261-327.

Higginbotham, J. (1985) "On Semantics." *Linguistic Inquiry* 16, 547-593.
Higginbotham, J. (1987) "Indefiniteness and Predication." E.J. Reuland; A.G.B. ter Meulen (eds.) (1987), 43-70.
Hirschbühler, P. (1975) "On the source of lefthand NPs in French." *Linguistic Inquiry* 6, 155-165.
Hoekstra, T. (1983) "The Distribution of Sentential Complements." In: H. Bennis; W.U.S. van Lessen Kloeke (eds.): *Linguistics in the Netherlands 1983*. Dordrecht: Foris, 93-103.
Horvath, J. (1986) *FOCUS in the theory of Grammar and the syntax of Hungarian*. Dordrecht: Foris.
Huang, C.T.J. (1982) "Move-wh in a language without wh-movement." *The Linguistic Review* 1, 369-416.
Huang, C.-T.J. (1984) "On the Distribution and Reference of Empty Pronouns." *Linguistic Inquiry* 15, 531-574.
van Hoof, H. (1994) Paper given at the Tilburg Workshop on Left Dislocation.
de Hoop, H. (1992) *Case Configuration and Noun Phrase Interpretation*. PhD diss., Groningen University
Iatridou, S. (1994) "On the Contribution of Conditional *then*." (to appear in: *Natural Language Semantics*.)
Iatridou, S. (forthcoming) "Clitics and Island Effects." to appear in: *Linguistic Inquiry*.
Jackendoff, R. (1977) *X'-Syntax. A Study of Phrase Structure*. Cambridge, Mass.: MIT Press.
Jacobson, P.; P. Neubauer (1976) "Rule Cyclicity: Evidence from the Intervention Constraint." *Linguistic Inquiry* 7, 429-461.
de Jong, F. (1987) "The compositional nature of (in)definiteness." E.J. Reuland; A.G.B. ter Meulen (1987) (eds.), 270-285.
Joseph, B. (1977) "On the Cyclicity of Extraposition-from-*the claim*." *Linguistic Inquiry* 8, 169-173.
Kadmon, N. (1990) "Uniqueness." *Linguistics and Philosophy*, 13, 273-324.
Kayne, R. (1975) *French Syntax: the transformational cycle*. Cambridge, Mass.: MIT Press.
Kayne, R. (1993) "The Antisymmetry of Syntax." ms. CUNY.
Kamp, H.; U. Reyle (1993) *From Discourse to Logic*. Dordrecht: Kluwer.
Kaplan, D. (1977) "Demonstratives. An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals." Published in: J. Almony; J. Perry; H. Wettstein (eds.) (1989) *Themes from Kaplan*. Oxford University Press, 481-563.
Karttunen, L. (1969) "Pronouns and variables." In: *CLS 5*.
Karttunen, L. (1974) "Presupposition and Linguistic Context." *Theoretical Linguistics* 1, 182-194.
Karttunen, L. (1976) "Discourse Referents." J. Mac Cawley (ed.) *Syntax & Semantics, vol. 7: Notes from the Linguistic Underground*. New York: Academic Press.
Karttunen L.; S. Peters (1979) "Conventional Implicature." C.-K. Oh; D.A.Dinneen (eds.) *Syntax and Semantics, vol. 11: Presupposition*. New York: Academic Press, 1-56.
Koopman, H. (1993) "The internal and external distribution of pronominal DPs" ms. UCLA, Los Angeles, California.
Koster, J. (1978) "Subject Sentences don't exist." In: S.J. Keyser (ed.) *Recent transformational studies in European languages*. Cambridge, Mass.: MIT Press.
Koster, J. (1981) *Locality Principles in Syntax*. Dordrecht: Foris.
Koster, J. (1987) *Domains and Dynasties*. Dordrecht: Foris.
Koster, J. (1993) "Predicate Incorporation and the Word Order of Dutch." ms. Groningen University
Kratzer A. (1989) "Stage-Level and Individual-Level Predicates." E. Bach; A. Kratzer; B.Partee (eds.) *Papers on Quantification*. Umass/Amherst.
Kripke, S. (1977) "Speaker's Reference and Semantic Reference." *Midwest Studies in Philosophy II*, 255-276.
Ladusaw, W. (1980) *Polarity Sensitivity as Inherent Scope Relations*. PhD. diss., University of Texas, Indiana University Press, Bloomington.
Ladusaw, W. (1980) "On the notion Affective in the analysis of negative polarity items." *Journal of Linguistic Research*, 1.

Raposo, E. (1986) "On the Null Object in European Portuguese." O. Jaeggli, C. Silva-Corválan (eds.) *Studies in Romance Linguistics*. Dordrecht: Foris. 373-390.
Reinhart, T. (1980) "On the position of extraposed Clauses." *Linguistic Inquiry* 11, 621-624.
Reinhart, T. (1983) "Coreference and bound anaphora." *Linguistics and Philosophy* 6, 47-88.
E. Reuland; A.G.B. ter Meulen (eds.) (1987) *The Representation of (In)definiteness*. Cambridge Mass.: MIT Press.
van Riemsdijk, H.C. (1978) *A Case Study in Syntactic Markedness. The binding nature of prepositional phrases*. Dordrecht: Foris.
van Riemsdijk, H.C.; E. Williams (1981) "NP-Structure." *The Linguistic Review* 1, 171-217.
van Riemsdijk, H.C.; E. Williams (1986) *Introduction to the Theory of Grammar*. Cambridge, Mass.: MIT Press.
van Riemsdijk, H.C.; F.Zwarts (1974) *Left Dislocation in Dutch and the status of copying rules*. unpublished ms., MIT/Univ. of Amsterdam.
van Riemsdijk, H.C.; F.Zwarts (1994) *Left Dislocation: 20 years later*. Handout for the Workshop on Left Dislocation, Tilburg.
Rizzi, L. (1986) "Null objects in Italian and the theory of *pro*." *Linguistic Inquiry* 17, 501-558.
Rizzi, L. (1990) "Speculations on Verb Second." J. Mascaró, M. Nespor (eds.) *Grammar in Progress. Glow Essays for Henk van Riemsdijk*. Dordrecht: Foris, 375-386.
Rizzi, L. (1990) *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rochemont, M. (1986) *Focus in Generative Grammar*. Amsterdam/Philadelphia: John Benjamins.
Rochemont, M.; P. Culicover (1990) *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.
Rooth, M. (1985) *Association with Focus*. PhD diss., GLSA, University of Massachussetts, Amherst.
Rooth, M. (1992) "A Theory of Focus Interpretation." *Natural Language Semantics* 1, 75-116.
Ross, J. (1967) *Constraints on variables in Syntax*. PhD diss., Cambridge, Mass.: MIT.
Rouveret, A. (1978) "Result Clauses and Conditions on Rules." S.J. Keyser (ed.) *Recent Transformational Studies in European Languages*. Cambridge, Mass.: MIT Press.
Russell, B. (1905) "On Denoting." *Mind* 14, 479-493.
Safir, K. (1985) "Missing Subjects in German." J.Toman (ed.) *Studies in German Grammar*. Dordrecht: Foris, 231-287.
Safir, K. (1986) "Relative clauses in a theory of binding and levels." *Linguistic Inquiry* 17, 663-690.
de Saussure, F. (1915) *Cours de Linguistique Generale*, reprinted as: Ch. Bally; A. Sechehaye (eds.) (1955). Paris: Payot.
Sperber, D.; D. Wilson (1986) *Relevance: communication and cognition*. Cambridge, Mass.: Harvard University Press.
Stalnaker, R. (1978) "Assertion." P. Cole (ed.) *Syntax and Semantics, vol. 9: Pragmatics*. New York: Academic Press.
Sportiche, D. (1983) *Structural Invariance and Symmetry in Syntax*. PhD. diss., Cambridge, Mass.: MIT.
Stowell, T. (1979) "Restrictive Relatives." E.Battistella (ed.) *Proceedings of the Ninth Annual Meeting of NELS. GLSA*. University of Massachusetts, Amherst.
Stowell, T. (1987) "Subjects, Specifiers and X-Bar Theory" ms. UCLA, Los Angeles, California.
Strawson, P.F. (1950) "On Referring." *Mind* 59, 320-344.
Thomason, R. (ed.) (1974) *Formal Philosophy. Selected Papers of Richard Montague*. New Haven and London: Yale University Press.
Trinker, B. (in preparation) *Satzkonstituenten im Deutschen Mittelfeld*. MA-thesis, Univ. Vienna.
Vallduví, E. (1990) *The Informational Component*. PhD diss., UPenn, Pennsylvania.
Vergnaud, J.R. (1974) *French Relative Clauses*. PhD diss., Cambridge Mass.: MIT.
Vergnaud J.R.; M.L. Zubizaretta (1990) "On the form and interpretation of definite noun phrases." unpublished ms. USC, Los Angeles.

Vergnaud J.R.; M.L. Zubizaretta (1992) "The Definite Determiner and the Inalienable Constructions in French and English." *Linguistic Inquiry* 11, 203-238.
Vikner, S. (1993) Verb Movement and Expletive Subjects in the Germanic Languages. (to be published by Kluwer).
Ward, G. (1985) *The semantics and pragmatics of preposing* University of Pennsylvania, PhD dissertation.
Ward, G.; B. Birner (1994) "English *there*-sentences and information status." R. Buchalla; A. Mittwoch (eds.) *The Israel Association for Theoretical Linguistics IATL 1 Proceedings of the Ninth Annual Conference Ben Gurion University of the Negev 1993 and of the Workshop on Discourse, Hebrew University of Jerusalem (1993)*, 165-183.
Wasow, T. (1972) *Anaphoric relations in English*. PhD diss., Cambridge, Mass.: MIT.
Williams, E. (1974) *Rule Ordering in Syntax*. PhD diss., Cambridge, Mass.: MIT.
Williams, E. (1977) "Discourse and Logical Form." *Linguistic Inquiry* 8, 101-139.
Williams, E. (1980) "Predication." *Linguistic Inquiry* 11, 203-238.
Williams, E. (1983) "Semantic vs. Syntactic Categories." *Linguistics and Philosophy* 6, 423-446.
Williams, E. (1986) "A Reassignment of the functions of LF." *Linguistic Inquiry* 17, 265-299.
Williams, E. (1993) *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.
Williams, E. (1994a) Talk given at the Vienna GLOW conference.
Williams (1994b) "Blocking and Anaphora." Ms. Princeton University.
Williams (1994c) "VP-ellipses." Ms. Princeton University.
Wiltschko, M. (1993) "Extraposition in German" *Wiener Linguistische Gazette* 48-50, 1-30.
Wiltschko, M. (1993) *ProPPs im Deutschen*. MA thesis, Univ. Wien.
Wiltschko, M. (in print) "Left Dislocation as Parasitic Operators." to appear in: *Proceedings of the Tilburg Workshop on Left Dislocation*. (April 1994).
Wiltschko, M. (in print) "Extraposition, Identification and Precedence", to appear in: *Proceedings of the Tilburg Conference on Rightward Movement* (October 1994)
Zubizaretta, M.-L. (1994) "The Grammatical Representation of Topic and Focus: implications for the Structure of the Clause." *University of Venice. Working Papers in Linguistic*, Vol.4.1, 97-126.
Zwart, J.-W. (1992) "SOV languages are head-initial" ms. University of Groningen.
Zwart, J.-W. (1993) *Dutch Syntax: A Minimalist Approach*. PhD. diss., University of Groningen.