1 Introduction

This study addresses the relative construction. It mainly concerns the syntax of relative clauses. Considerable attention is also paid to the typology of relativization, and its repercussions on the grammar. Section 1 below introduces the topic of relativization; section 2 delineates the objectives and contents of this thesis; and section 3 discusses the syntactic framework used.

1. Preamble: the problem of the pivot

Relative constructions have received much attention by linguists, and not without reason. They are highly interesting from a syntactic, typological, and semantic point of view. As an introduction consider the ‘problem of the pivot’.

What distinguishes relative clauses from other subordinate clauses is that there is a direct link between an element in the relative and in the matrix. To put it even stronger: there is a pivot element that plays a role in both. For instance, in (1) this pivot is books.

(1) Jack never reads *books* I recommend *to* him.

It may be problematic to put it this way. How can *books* have two functions at the same time? And what is the position of the relative clause in the matrix? A potential solution is to assume the representation in (2), where an empty element that is coreferent with *books* occupies the direct object position within the relative clause. This makes *books* an antecedent, which takes the position of the direct object in the matrix. The relative is somehow attached to this antecedent.

(2) Jack never reads *books*, *I recommend* it *to* him.

The existence of relative pronouns may justify such a move. A relative pronoun may be the overt counterpart of the empty element. If a relative pronoun is used, it must be fronted within the relative. This suggests that it is raised from the object position, as an interrogative pronoun does in questions. See (3), where the base position of *which* is indicated with the trace t.

(3) Jack never reads *books*, [which, *I recommend* t, to him].

If the empty element in (2) is equivalent with the relative pronoun, it could be that it is raised, too. The adjacency reached may facilitate the link between *books* and ø/which; see (4).

(4) Jack never reads *books*, [ø, *I recommend* t, to him].
But if there is raising in (4), why would the empty element be raised, and not simply *books* itself, as in (5a)? Subsequently, it could even be lifted to the main clause; as shown in (5b).

(5)  
a. Jack never reads *[books, I recommend it to him]*.  
b. Jack never reads *books, [it, I recommend it to him]*.

Alternatively, both could be raised together; see (6).

(6)  
a. Jack never reads *[books of, I recommend it to him]*.  
b. Jack never reads *[books which, I recommend it to him]*.

Finally, it is possible to approach the problem from the opposite side and assume that it could be the empty element that is in the matrix, and the noun *books* that is in the relative; see (7).

(7)  
Jack never reads *a, [[books (which), I recommend it to him]*.

This naïve exposition shows that there are several strategies to cope with the 'problem of the pivot'. Therefore it is not surprising that languages have found different but related ways to express the relative construction. A cross-linguistic typological survey indicates that there are four syntactic main types: the postnominal, prenominal, and circumnominal relative, and the correlative. The literal equivalents of these in English are illustrated in (8a) through (8d), respectively. The intended meaning is the same in each case.

(8)  
a. Jack never reads *books (which) I recommend to him*.  
b. < Jack never reads *[I recommend to him] books.* >  
c. < Jack never reads *[I recommend books to him]*. >  
d. < *[Which books I recommend to him] Jack never reads them.* >

Hence the position where the pivot is pronounced, is variable across languages. Furthermore, there is variation concerning the use of relative elements, as indicated. English itself has three out of four obvious variants:

(9)  
Jack never reads *books of/which/that I recommend (*them) to him.*

The first is called 'zero relativization'. Second, a relative pronoun *which* can fill the 'gap' in the relative. Third, the relative clause can be introduced by a relative complementizer *that*. Fourth, some languages use a resumptive pronoun. This variation can be used to determine the position and function of potential empty elements.

Thus there is a clear interaction between the syntax and typology of relative clauses in the sense that the typological variation offers invaluable information to determine the right analysis, and, on the other hand, syntax has the ability to explain
– or at least describe – the attested variation, and to show the relation between the different constructions.

2. Goals and contents of this thesis

Here I will first set out the primary objectives of this thesis in section 2.1. Next, section 2.2 provides an overview of the contents of the separate chapters.

2.1. Objectives

Acknowledging the danger of passing over important work, I believe the following studies can be viewed as major breakthroughs in the conception of the relative construction:

– Chomsky (1977), regarding the internal syntax of relative clauses;
– Lehmann (1984), regarding the typology of relative clauses;
– Grosu & Landman (1998), regarding the semantics of relative clauses.

In practice there turn out to be several types of relatives, both syntactically and semantically. In this context, I must mention in particular:

– Carlson (1977), regarding the syntax and semantics of degree relatives;
– Culy (1990), regarding the syntax and semantics of circumnominal relatives;
– Srivastav (1991), regarding the syntax and semantics of correlatives.

For an explanation of the relevant notions, see Chapter 2.

Other important references are the paper collection in Peranteau et al. (1972), which focuses on relative elements; Smits (1988), an overview volume with detailed descriptions of Germanic and Romance relatives; Bianchi (1999), regarding the promotion theory of relative clauses; and the paper collection in Alexiadou et al. (2000). Apart from these, there is a large amount of articles on relative constructions by various authors.

At present there is still a debate in generative syntax between the proponents of the ‘standard analysis’ and those of the ‘raising analysis’. These are illustrated in a simplified form in (10a/b), respectively.

(10) a. Jack never reads books, [a, I recommend t, to him].
   b. Jack never reads [books, I recommend t, to him].

The major difference is that the head noun books is generated in the matrix clause in (10a), but it is raised from within the relative in (10b). The raising or ‘promotion’ analysis has originally been proposed by Schachter (1973) and Vergnaud (1974, 1985). It has been revived and modernized by Kayne (1994) within a general antisymmetric framework of syntax. In fact, Kayne combines raising with the D-complement hypothesis, which states that the relative clause is the complement of
the outer determiner. The promotion theory is worked out in more detail by Bianchi (1995,1999) and De Vries (1996). Matters that are addressed are e.g. the categorial status of the projections involved, the position of relative elements, the way Case is accounted for, and the triggers for the movements involved.

Kayne (1994) is the first who seriously tries to generalize over several relative clause types, although the discussion is still rudimentary. I believe that this is the right track to follow. It seems that it is made possible by the characteristics of the promotion theory. Therefore a priori this theory appears to be preferable over the variants of the standard analysis. However, in the present study – cf. De Vries (1997) for an early description – I will not simply take it for granted and focus on the details, nor will I follow the antisymmetry hypothesis without discussion.

In short, the primary objectives of this thesis are the following (in order of treatment):

- to provide a systematic and accessible typology of relative constructions as a background for this and future inquiries;
- to compare extensively the competing syntactic approaches to relativization;
- to provide a detailed and consistent account of the syntax of various types of relative constructions;
- to provide an independently motivated solution to potential problems not satisfactorily addressed before (from the perspective of the promotion theory); these are:
  - the syntax of appositive relative clauses,
  - extraposition of relative clauses,
  - the syntax of possessive relatives.

The next subsection discusses briefly the contents of the separate chapters of this book.

2.2. Overview

The first part of the book, which consists of four chapters (2 through 5), discusses the typology and the theory of relative constructions.

Chapter 2 provides a typology of relative constructions. It defines and exemplifies the main types of relative clauses, shows the parametric freedom there is, systematizes classifications, and unifies the terminology. It discusses Grosu & Landman's semantic classification of relatives, Lehmann’s functional scales, and Downing’s universals and general implications. It introduces special types of relatives, viz. correlatives, circumnominal relatives, free relatives, adverbial relatives, and non-finite relatives; and the cleft and pseudo-cleft construction. Finally, several aspects of the relative construction are treated briefly: relative pronouns and particles, the position of the external determiner, recursive and linear multiple embedding (including stacking), pied piping and preposition stranding, extraposition, and multiple relativization (including split antecedents).

Chapter 3 works its way towards the syntax of relative clauses. It sketches the historical development of the theory on the syntax of relativization, and explains and
INTRODUCTION

defends the D-complement hypothesis and the raising analysis. The most important competing theories are defined and systematically evaluated on the basis of possible derivations of the syntactic main types of relatives and their word order variants, and on the basis of relevant properties of relative constructions, mainly concerning the relation between the antecedent and the gap. It is concluded that the promotion theory is the most promising.

Chapter 4 elaborates the promotion theory in detail. It discusses *wh*-movement, Case and agreement, the relation between N and D, the role of relative pronouns, and triggers for movements. The derivation of word order variants of the postnominal relative construction is discussed, as well as the derivation of prenominal relatives, circumnominal relatives and correlatives.

Chapter 5 is on relative elements. Lehmann’s classification of these is revised on the basis of the syntax of relatives. Somewhat tentatively an analysis of resumptive pronouns and relative markers is proposed. Finally, a fine-grained typology of relative elements is presented.

The second part of the book (Chapters 6 through 8) contains three related detailed studies which focus on relative constructions, but also have a more general character.

Chapter 6 discusses the syntax of apposition, and of appositive relatives in particular. It contains a large collection of properties of appositive relatives, especially those in which they deviate from restrictives. It is argued that apposition is specifying coordination. This leads to an analysis in which appositive relatives are so-called false free relatives that are specifying conjuncts to the ‘antecedent’. In this conjunct an empty pronoun is promoted.

Chapter 7 treats extraposition, in particular of relative clauses. It systematically evaluates a number of competing theories on the basis of a substantial amount of properties associated with extraposition. It is concluded that extraposed phrases are part of a specifying conjunct to the matrix, in which deletion takes place. This approach can be generalized to all instances of extraposition. Crucially, the promotion theory of relatives can be maintained, since promotion is performed within this second conjunct.

Chapter 8, finally, addresses possessive constructions, and possessive relatives in particular. It is argued that all possessive configurations are syntactically derived from the periphrastic genitive. Given this framework, promotion in possessive relatives can be analysed along the lines of pied piping with prepositional phrases. Eventually, cases of heavy pied piping are discussed.
3. **The theory of grammar**

This section briefly describes the general syntactic framework within which this thesis must be understood. The way I implement specific ideas (i.e. about relative constructions) does not differ substantially from general practice these days, which can be characterized as 'inclined to the Minimalist Program, with a flavour of Antisymmetry'. Nevertheless, I would like to add some details and critical remarks here.

### 3.1. Phrase structure

Syntactic phrase structures are derived by concatenating elements from the lexicon. For non-linguists this sounds like stating the obvious. Therefore it is remarkable, to say the least, that it lasted until the nineties before the idea was implemented in a direct way in generative linguistics. In principle, there are two ways to proceed: top-down and bottom-up. If the procedure is top-down, the equivalent of a D-structure is derived first. After lexical insertion, the necessary movements must be performed, hence the derivation turns around and works its way from the bottom to the top again (now involving movement, not structure building). This gives an S-structure representation. This procedure is somewhat laborious, and indeed, it can be done in a more clever way, namely if the procedure is bottom-up from the beginning. This is also the standpoint of the Minimalist Program (Chomsky 1995), in which lexical insertion and movement are structure building. Another difference with the top-down method is that the strict separation of selection from the lexicon and movement is lost, in the sense that both are performed interchangeably during the derivation. Depending on the particular feature setting and the intended meaning, one of three possible elements is merged with the phrase structure at a particular point of the derivation: an element from the lexicon, a moved element from the phrase structure itself, or another partial phrase structure.

At this point I want to introduce the theory-external notion *selection structure*, which is used for notation only. The selection structure of a (partial) sentence structure is the hypothetical equivalent of a D-structure, i.e. a structure that shows all Merge operations that access the lexicon, but not those that involve movement. Obviously this has no reality at all within the derivational theory described, but it is still useful to show what selects what.

Following general practice, I assume that the phrase structure is binary branching. In general, it can be demonstrated that for any pair of constituents in a sentence either there is a hierarchical relation between them, or they are coordinated. Binary branching is encoded in the formulation of the operation

---

1. It is shown in Chapter 7 that a binary branching phrase structure cannot handle coordination in a satisfactory way. Therefore a three-dimensional approach is developed, following Van Riemsdijk (1998) and others. In Chapters 6 and 7 I will also introduce the concept of specifying coordination as a basis of the analysis of both apposition and extraposition.
Merge. Furthermore, the phrase structure is organized according to the recursive $X'$-schema: \([xPZ[x XYP]]\), where the linear order between sister constituents might be parametrically interchangeable (but see below). The complement of the head X is YP; the specifier is ZP. The categorial status of X(P), Y(P) and Z(P) follows from their feature setting. For the ease of representation, I will name particular projections according to the setting of their heads throughout this book. Adjuncts, if possible at all, are attached at the highest XP level.

I have no objections to encoding the simple $X'$-template directly in the grammar, but there are several attempts in the literature to derive it, or its most important properties, from more basic assumptions; see e.g. Chomsky (1995:Ch4). Of course, if tenable, this is to be preferred. These matters do not directly concern the subject of this book, therefore I will not discuss them any further, with one exception: Kayne's (1994) argument concerning the "Antisymmetry of syntax". He claims that the order of spec-head-comp is fixed, and that this follows from the Linear Correspondence Axiom, which involves the mapping of the syntactic phrase structure hierarchy onto a linear order. In other words: the fact that sentences must be linearized at or beyond the phonological interface imposes restrictions on the syntactic hierarchy.\(^2\) (Of course this is only true from the perspective of Kayne's premises, since a tree scanning algorithm can linearize any tree, including those that are not antisymmetric; see below.) Kayne suggests that the LCA is part of Universal Grammar. He is not very clear about the way it is encoded in the grammar. To me, the LCA looks like a filter, hence there could be an 'LCA checking procedure' that filters out phrase structures that are wrong (i.e. not linearizable). If so, we find that, although the LCA is a beautiful theory, it is not very efficient, compared to a primitive $X'$-template in combination with a simple tree scanning algorithm,\(^3\) since the complexity of a scanning procedure grows linearly with the size of a phrase structure, whereas that of an LCA checking procedure grows exponentially.\(^4\)

---

2 Notice that, strictly speaking, the tree structure \([\text{spec [head comp]}]\) and the structures with the same hierarchy \([- [\text{spec [comp head]}], [[\text{head comp} \text{ spec}]] \text{ and } [[\text{comp head} \text{ spec}]]\) are all linearly projected as spec-head-comp. Notationally, only the first is workable, hence I will use that one.

3 An example of such an algorithm is the following:

```
start at top; create new string
→ if there exists an unexplored node one step down left go there
   else if there exists an unexplored node one step down right go there
   else if possible go one step up
   else stop
mark present node as explored
if terminal then add lexical material to string
loop →
```

4 If a certain small phrase structure can be linearized in $x$ steps, where $x$ depends on the algorithm and the number of terminals and non-terminals, then a larger structure which is about $y$ times as big can be scanned in roughly $xy$ steps. By contrast, if it takes $z$ steps to check if the small phrase structure can be mapped on a linear order according to the LCA, it takes roughly $zy^3$ steps to perform the same procedure on the larger structure. Consider for instance the transitivity property of a linear order: $\forall a \forall b \forall c: (a,b)$ and $(b,c)$ then $(a,c)$. Here $a$, $b$ and $c$ are elements from the set of terminals, and the relation is precedence. The checking of this condition requires a triple loop over all elements. Furthermore, the LCA checking procedure requires to establish the set of all pairs of non-terminals \(<X,Y>\) that are related by asymmetrical c-command. Given that c-command (of $Y$ by $X$) to be continued...
Therefore a filter version of the LCA probably cannot have any neuro-psychological reality. This does not mean that syntax is not antisymmetric. It may very well be that the X'-schema is rigidly fixed as [spec [head comp]], which is a rule of thumb translation of the LCA that can be accessed during the derivation. (This would make the original LCA to an extra-grammatical theory.) In Chapters 3 and 4 I will show on an empirical basis that (most probably) at least specifiers and functional heads are on the left. I will remain agnostic concerning full antisymmetry.

With respect to functional projections I will take the conservative standpoint that they cannot be used without extensive motivation. For my purposes I do not need any other projections than CP, IP, AgrOP and DP (next to the lexical NP, AP, VP and PP), even though quite complicated constructions will be dealt with. Of course this does not mean that I reject other possible projections across-the-board.

3.2. Movement and features

A head consists of at least three things: i) syntactic features, ii) semantic features and/or a pointer to the relevant semantic part in the lexicon, iii) phonological features and/or a pointer to the relevant part in the lexicon. Syntactic movement is driven by the need for syntactic feature checking. Unchecked features lead to a crash of the derivation at the conceptual/intentional interface (or LF). By assumption, there are two ways of checking a feature: i) in a spec-head configuration; ii) in a head incorporation configuration. I will assume that features that are checked simply receive a check mark.\(^5\) Since incorporation signifies the merger of two heads, there may not be contradictory features, and checking is obligatory. By contrast, I assume that in a spec-head configuration checking is neither obligatory, nor necessarily complete (in the sense that all relevant features that could be checked are checked).\(^6\) Obviously the possibilities of (temporary) incomplete checking are severely limited by the constraint that eventually all features must have a check mark. Notice that incomplete checking is necessary for intermediate landing positions.

\(^5\) Hence if necessary, they are accessible again. I do not favour a deletion mechanism, an erasure mechanism, and the rather stipulative difference between interpretable and uninterpretable features as in Chomsky (1995).

\(^6\) Hence I will not use the term spec-head agreement, which would be misleading. I will show that the contrast between spec-head and incorporation is useful. Notice, furthermore, that it would require an additional assumption to force complete checking in a spec-head configuration (which is assumed by many authors), rather than the opposite.
An important way to encode differences between languages is the distinction between strong and weak features. By assumption, strong features require checking by overt movement; weak features may be checked covertly. The latter is explained by an important economy condition, which states that overt movement is more expensive than covert movement. Covert movement implies that the phonological features/pointer are left behind, hence it is ‘lighter’.7

A derivational syntax implies strict cyclicity for structure building operations, hence for Merge and Move (which implies Merge). However, most people assume that LF movement (covert movement) is possible, which is completely at odds with the spirit of the approach.8 Instead I propose to give the ‘strict cycle condition’ (which is a simple consequence of Merge) a central place in the syntactic theory. Covert movement can then be viewed as ‘overt’ feature movement, or ‘partial head movement’, an option that is present in the theory anyway. This has many consequences. In general, it makes the theory less complicated. I will tentatively mention some of the issues at stake, although they require an elaborate discussion each.

First, the strange T-model of grammar can be abandoned, in which the derivation proceeds after spell-out/S-structure. Instead, the completed derivation is sent to the phonological and semantic interfaces. Second, since reconstruction at LF is impossible, it is necessary that at least some of the interpretation is done during the derivation. Especially binding facts come to mind. I have argued in De Vries (1998a) that indeed it is preferable that the binding conditions apply during the derivation.9 In fact, I believe that the semantic derivation proceeds parallel with the syntactic derivation. This can be accomplished, for instance, by feeding the semantic component at certain points of the derivation, e.g. after the completion of each predication (i.e. a maximal (extended) lexical projection).10 If all this is correct, there may be no use in keeping track of the origin of moved constituents in syntax, hence the concept of trace (or copy), and perhaps even chain, can potentially be eliminated from the theory.11,12 I will not discuss these matters any further here.

Finally, I assume that features are associated with heads. This, and the concept of overt feature movement offers an explanation for the phenomenon of pied piping. First, notice that the features of a head must be accessible to the level of the maximal projection. This is sometimes called ‘feature percolation’, and in fact it is the most elementary form of pied piping. If an XP moves to the specifier of Y, this may be viewed as pied piping of the XP, as an alternative to potential head

---

7 Since economy preferences are overridden if the cheaper option leads to a crash of the derivation, it is predicted that it is possible that a weak feature must be checked overtly in exceptional circumstances. An example of this is discussed in Chapter 4.
8 Notice that LF-movement implies going over the entire derivation a second time. As I see it, this means throwing away a major advantage over pre-Minimalist approaches.
9 Others have made comparable claims. See e.g. Zwart (1999).
10 A similar idea has been proposed by Chomsky (1999).
11 For instance, the fact that a trace behaves as an anaphor can be derived from other notions, such as a ‘shortest step condition of movement’ and the fact that movement is always to a c-commanding position due to the character of Merge.
12 Nevertheless, for the ease of representation traces are indicated throughout this book.
movement of X to Y. XP and Y are in a spec-head configuration, then, so there is checking between some features of XP and Y. ‘True’ pied piping can be the result of percolation (i.e. overt feature movement) to a higher head. For instance, the wh-feature of a nominal phrase may move to a higher prepositional head, which causes pied piping of the whole PP to SpecCP. Percolation is only possible if the target does not bear a similar kind of feature; see Chapter 4. Another possible cause of pied piping is head movement. This serves to check some feature(s), but some other feature(s) of the moved head may still be unchecked. In turn this would cause movement, which results in pied piping (either of a complex head or a maximal projection). See further Chapter 8 on possessive relatives.

\[13\] Therefore it may be that there is no primitive distinction between X and XP features. (The landing position for XP movement is different from that of head movement, however.) Notice that many potential head movements are impossible. For instance, the head D of a subject argument cannot move to I, since there is no word that expresses a determinerhood with temporal inflection. Put more generally: incorporation is impossible if there are contradictory features, as noted before. Thus, in this case, the subject DP moves to SpecIP where some features are checked (here: number and person), but not those that are contradictory (e.g. +/− V).