The syntax of relativization

de Vries, M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
4 The promotion theory of relative constructions

1. Introduction

Taking the conclusions of Chapter 3 as a starting point, this chapter discusses the syntax of (restrictive) relative clause constructions in terms of the promotion theory of relativization in detail. Section 2 is an introduction to the promotion theory, and an outline of earlier work by Kayne (1994) and Bianchi (1999), upon which this chapter is based in part. The core of my proposal concerning postnominal restrictive relatives is outlined in section 3. Sections 4 through 6 explain how it can be extended to the other syntactic main types of relatives: prenominal, circumnominal and correlative. Relative pronouns and particles are discussed separately in the next chapter. Section 7 summarizes and concludes the chapter.

2. The promotion theory: previous scholarship

The promotion theory has its roots in Schachter (1973) and Vergnaud (1974/1985). As discussed in Ch3§3.1.3, the raising analysis as proposed by these authors suffers from serious problems from the perspective of present-day syntax. Subsequently, in Ch3§3.2/3, I have shown that even the revised raising analysis performs less well than the promotion theory, which combines raising with the D-complement hypothesis. Therefore I will take Kayne's (1994) original proposal of the promotion theory as a starting point here.

2.1. Kayne’s (1994) analysis and Borsley’s (1997) criticism

In Kayne's system a relative construction like *the house (that) I painted* is represented as in (1). The relative CP is selected by the outer determiner *the*, and the antecedent *house* by the subordinate verb *painted*. The selection structure is shown in (1a). The antecedent *house* is moved to SpecCP; see (1b).

\[
\begin{align*}
(1) & \quad \text{a. } [\text{DP} [D' \text{ the } [\text{CP} (\text{that}) \text{ I painted house }]]] \\
& \quad \text{b. } [\text{DP} [D' \text{ the } [\text{CP} \text{ house} [C' (\text{that}) \text{ I painted } t]]]]
\end{align*}
\]

If there is a relative pronoun (Drel), e.g. in *the house which I painted*, the situation is a little more complex. The pronoun originally takes the position that normally a determiner does: \[D: [w \text{ which}] [\text{NP house}].\] It is this complex that raises; see (2b). But the word order is still not the final one: the head *house* needs to precede *which*. Therefore NP moves to SpecDP, as shown in (2c).
On may wonder why the relative pronoun is not simply put in the position of C, as in (3).

(3)  [DP [D' the [CP [NP house], [C which [IP I painted t]],]]]

But this is impossible. Several languages show both a relative pronoun and a complementizer (cf. Chapter 5, sections 3.1 and 4.2). For instance, (4a) is translated Middle English, (4b) is an example from a dialect of Dutch (i.e. Aarschot), taken from Dekkers (1999:58).

(4)  a.  the man who that I saw
    b.  de stoelen di da kapot zijn

So the complementizer position is already occupied. Also see Lehmann (1984), Hoekstra (1994), Pitter (1996), Bianchi (1999), Broekhuis & Dekkers (2000), and others for data on this matter.

Could not a relative pronoun have its own projection, then? This idea is sketched in (5), where the antecedent and the relative CP are the specifier and the complement of the relative pronoun head, respectively.

(5)  [RelP antecedent [Rel who [relative CP]]]

The suggestion in (5) can have many theoretical variants, with or without raising. All are clearly wrong, for several reasons. First, in Lehmann's terms (cf. Ch2§4), one of the functions of a relative pronoun is *Gap Construction* (German: 'Leerstellenbildung'). In many languages this can be detected easily, since a relative pronoun bears subordinate clause Case. This is shown by the German example in (6).

(6)  Ich sah den Herrn der einen Hut trug.

Hence the Case on the pronoun shows that it is part of the relative clause. Therefore (5) cannot be the basic structure of a relative construction. Second, the relative pronoun can be part of a larger pied piped constituent, e.g. *whose mother*, or *in
which. Hence the wh-phrase cannot be a head. It must be a maximal projection which is moved to the front of the relative clause: SpecCP.

Thus consider again (1b) and (2c), repeated in (7a/b). In Kayne's original proposal there is no DP_{rel} projection in (7a). Raising applies just to NP (or QP if there is an additional quantifier).

(7) a. \[ \text{DP} \left[ D^* \text{the} \left[ \text{CP} \left[ \text{NP} \text{house} \right] \right] \left[ C^- \text{(that I painted)} \right] \right] \]
   
   b. \[ \text{DP} \left[ D^* \text{the} \left[ \text{CP} \left[ \text{DP}_{rel} \left[ \text{NP} \text{house} \right] \right] D^-_{rel} \text{which} t_i \right] \right] \left[ C^- \text{I painted} t_i \right] \]

It is bothersome that the derivation in (7a) differs from (7b). In De Vries (1996) I have argued that the two sentences have the same structural analysis: (7b). The only difference is that the first sentence has a zero relative pronoun. This is confirmed in a sense by Borsley (1997). He shows that the gap position in the relative clause is a DP position for several reasons. First, the gap is an argument position, and arguments are DPs (cf. Abney 1987, Longobardi 1994 and others). Second, the trace acts as a variable, hence as a DP-trace, with respect to several tests: binding, control, licencing of parasitic gaps, Case marking and weak islands. This is illustrated below, where I use Borsley's (1997:632/3) data.

In (8), the trace of a non-wh-relative can be co-indexed with a pronoun, provided that the pronoun does not c-command it. This parallels the data in (9) with DP-traces in wh-questions.

(8) a. the man, that t_i thought he, saw a UFO
   b. the man, that he* thought t_i saw a UFO

(9) a. Who, t_i thought he, saw a UFO?
   b. Who, did he* think t_i saw a UFO?

In (10) the trace controls a PRO subject.

(10) a. the man, that t_i tried PRO, to fool everybody.
   b. Who, t_i tried PRO, to fool everybody?

In (11) it licences a parasitic gap.

(11) a. the book, that Bill criticized t_i, without reading e_i
   b. Which book, did Bill criticize t_i, without reading e_i?

The trace (or chain of traces) must occupy a Case-marked position, hence expletive constructions like (12) are excluded. Again, this is similar to the situation in wh-questions as in (13).

(12) a. the man, that (*it) was arrested t_i
   b. the man, that (*it) seemed t_i to know the answer
(13) a. Which man, was (*it) arrested t₁?
   b. Which man, {seemed}/* {did it seem} t₁ to know the answer?

Finally, some languages allow extraction of referential phrases from weak islands. Extraction is also possible in non-wh-relatives; see (14).

(14) a. the book, that we wondered how to afford t₁ [infinitival wh-compl]
   b. the book, that we regretted that John read t₁ [factive complement]

If it is assumed that referential phrases are DPs, the gap in a relative clause is probably also a DP.

Thus I conclude with Borsley that the gap in a relative clause must be a DP, whether there is an overt relative pronoun or not. This is confirmed by Bianchi (2000a).

Apart from the issue concerning the categorial status of the gap, questions like the following must be answered:

- What forces the movements indicated in (7b)?
- How can the Case patterns in relative constructions be explained?
- Why can a determiner head be empty in a relative DP in a that-relative, but not elsewhere?

Kayne (1994) hardly addresses these matters. Borsley (1997) assumes that the promotion theory needs many additional ad hoc mechanisms to get things right, and therefore he rejects it altogether. However, his critique was anticipated upon and partly countered in Bianchi (1995) and De Vries (1996). Moreover, in Ch3§3 of this book I have shown that the standard theory actually needs more additional mechanisms. The second part of Borsley's critique concerns apposition and extraposition. These subjects are treated separately in Chapters 6 and 7, respectively, where I show them to be independent of the specific theory of relativization chosen.

2.2. Bianchi (1999/2000a)

At this point, consider Bianchi's (1999/2000a) revision of Kayne (1994). She also distinguishes that-relatives from wh-relatives, but in another way. The relevant structures are given in (15), cf. Bianchi (2000a:125/130).

(15) a. [DP Drel+the [CP [DP tD-rel picture], [C that Bill liked t₁]]]
   b. [DP the [CP [NP picture] [C that Bill liked t₁]] [XP which t₁p], [X’ X [IP Bill liked t₁]]]]

Bianchi argues the following:

(i) The antecedent is raised to SpecCP, because the outer D has a strong selectional feature that can only be checked by a [+N] category in its minimal domain.
(ii) The empty D_rel in that-relatives is licenced by incorporation into the external D.
(iii) *Wh*-relatives have a split CP à la Rizzi (1997). DP\textsubscript{rel} is only raised to SpecXP; NP moves on to SpecCP (instead of SpecDP\textsubscript{rel}, as Kayne assumed).

Although Bianchi's analysis has its advantages, I will not follow her exactly on these points.

In my view, selection cannot be feature checking as Bianchi presents it. In fact, Bianchi seems to introduce a third type of feature (a 'subcategorization feature'?), which combines traditional selectional features of the lexicon with formal syntactic features that drive movement. But there are also practical problems with respect to (i). In (15b), for instance, D's selectional feature [+N] must checked eventually by the raised noun phrase. At a certain point of the derivation, D and CP are merged, but at that very moment no checking can be performed, since NP is still in SpecXP. Thus NP must be raised to SpecCP first. This could not have happened before the merging of CP with D, since there is no trigger for it within CP. However, overt movement of NP to SpecCP after merging CP with D is countercyclic movement. Moreover, it is not clear what has caused the presence of the CP level to begin with. These problems do not occur in my analysis, as will be explained in the next section.

Concerning (ii), the licensing of D\textsubscript{rel} by incorporating it into the external D is only legitimate if their features are compatible. Their \*-features match, so that is all right. Furthermore, Bianchi argues that D\textsubscript{rel} is underspecified for definiteness, hence it cannot disagree with D's feature specification in this respect. The problem may be the Case feature. Whether it is empty or not, D\textsubscript{rel} gets Case in the subordinate clause. Since the external D checks its Case in the matrix clause, the two Case features are generally incompatible (unless accidentally). According to Bianchi, D\textsubscript{rel}'s Case feature is already checked and erased before D\textsubscript{rel} is raised. Unfortunately this procedure is incompatible with the general view on features presented in Chapter 1 (where I claim that there is no 'erasure' of features).

As for (iii), I consider it particularly unattractive that *wh*-relatives and *that*-relatives get different analyses. For instance, why do *wh*-relatives have a split CP and *that*-relatives not? Moreover, (iii) raises other problems. Sentences like (4) above with both a complementizer and a relative pronoun show that D\textsubscript{rel} must be in the highest projection. The representation in (15b) leads to wrong word orders, viz. those where C would precede D\textsubscript{rel}, which is never attested.\(^2\) Second, if NP can move alone, what forces pied piping of DP\textsubscript{rel} (which is necessary for the raising of D\textsubscript{rel} later on) in (15a) except a backtracking procedure? Finally (but that may be a matter of execution), we need to know what forces movement of DP\textsubscript{rel} to SpecXP in (15b), which implies a precise answer to the question what XP actually is.

I will not discuss Bianchi's analysis at length. Clearly, it is far more elaborate than Kayne's original proposal and obviates a substantial part of Borsley's critique. Nevertheless, it is not completely compatible with the general assumptions in this book. In the next section I will present an alternative to her approach, the basis of

\(^2\) There may be one exception, though. Hoekstra (1994) reports the order of die 'if who' in the Amsterdam dialect of Dutch. However, notice that of is the wrong complementizer in this context. It is normally used in questions; in relative clauses we expect dat 'that'. This makes these data suspect.
which was laid in De Vries (1996). Subsequently, I will apply it to a large range of relative constructions in sections 4 through 6.

3. Postnominal relatives

This section treats of the basic syntax of postnominal relative constructions, the most common type of relativization. The derivation is presented in terms of feature checking. The rationale behind it is based on wh, Case and agreement facts. In order to be clear about the necessary assumptions, I have split up the analysis in small parts. First, section 3.1 deals with wh-movement; 3.2 explains the relevance of Case and agreement; 3.3 discusses the relation between D and N; 3.4 shows a detailed derivation of a canonical postnominal restrictive relative; 3.5 is on that-relatives; 3.6. discusses word order variation; 3.7 concludes the argument.

3.1. Wh-movement

Many relative pronouns are morphologically identical to interrogative pronouns (at least in the Indo-European languages; cf. Appendix II, table 8). Examples are which and who in English. Therefore the assumption that there may be wh-movement in relative clauses is plausible. In fact, since Chomsky (1977) it is generally accepted that relative clauses involve wh-movement. The diagnostics for it are listed in (16), quoted from Chomsky (1977:86).

(16) Wh-movement
   a. It leaves a gap.
   b. Where there is a bridge, there is an apparent violation of Subjacency, the Propositional-Island Condition (PIC), and the Specified Subject Condition (SSC).
   c. It observes the Complex Noun Phrase Constraint (CNPC).
   d. It observes wh-island constraints.

---

3 The definitions of the conditions in (16b/c) are stated below. I will not illustrate them separately here; see e.g. Chomsky (1977), Bach (1977), Ross (1967) and Van Riemsdijk & Williams (1986) for discussion.

Subjacency, taken from Chomsky (1977:73): “a cyclic rule [i.e. move NP/wh] cannot move a phrase Y to X (or conversely) in the structure ... X ... [a ... [α ... Y ...]β ...]α ... X ... where α, β are cyclic nodes [i.e. S’, NP].”

PIC, taken from Bach (1977:145): “given a structure ... X ... [α ... Y ...] ... X ... where α is a cyclic node (S’, NP possibly S, S”), no rule may involve X and Y if α is a propositional island (for English = finite clause).” [Note that this is a parametrized version of the so-called Tensed-S Condition.]

SSC, taken from Van Riemsdijk & Williams (1986:118): “No rule may relate X and Y in the structure ... X ... [α ... Z ... W_1 Y W_2 ...]α ... (or ... [α ... Z ... W_1 Y W_2 ...]α ...) where Z is the subject of W_1 Y W_2.”

CNPC, taken from Ross (1967:76): “No element contained in an S dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.”
What is particularly interesting is that *that*-relatives satisfy the criteria as well, although there is no overt *wh*-word. Below, all elements of (16a-d) will be illustrated, for both types: *wh*-relatives and *that*-relatives. The examples are mine. Whenever OP is used, this designates a moved *wh*-operator: the empty counterpart of a relative pronoun.\(^4\)

Property (16a) is rather obvious:

(17) a. the meal which, you ate \(t\); /*beans
   b. the meal OP, (that) you ate \(t\); /*beans

Since *which* is the moved object of *ate*, there cannot be another object.

Diagnostic (16b) means that if the movement seems unbounded (i.e. crossing sentence boundaries in one swoop) it involves successive cyclic movement via the COMP position (SpecCP). This possibility is illustrated in (18).

(18) a. the meal which, you predicted \(t\); "that Luke believed \(t\);" that Nana ate \(t\)
   b. the meal OP, (that) you predicted \(t"); "that Luke believed \(t\);" that Nana ate \(t\)

Constructions like (18) are marked or even impossible in some languages. In fact, they are often marked in English as well. What is important, though, is that the acceptability judgements exactly parallel those for parallel *wh*-question sentences, e.g. *Which meal did you predict that Luke believed that Nana ate?*

The effect of the Complex NP Constraint is shown in (19).

(19) a. *the meal which, I heard the story that Luke ate \(t\);
   b. *the meal OP, (that) I heard the story that Luke ate \(t\);

Finally, a *wh*-island construction is given in (20).

(20) a. *the meal which, I wondered who ate \(t\);
   b. *the meal OP, (that) I wondered who ate \(t\);

Here *who* is in SpecCP, hence it occupies the necessary bridge for movement of *which/OP*, thus causing a Subjacency violation.

Clearly, both types of relative clauses (*that*-relatives and *wh*-relatives) meet the criteria for *wh*-movement.\(^5\) In the present framework of syntax, *wh*-movement may be seen as movement of a constituent that bears a *wh*-feature to SpecCP, where the feature is checked. Given that there are empty operators that perform *wh*-movement, a *wh*-feature does not need morphological support: it has a more abstract nature.

---

\(^4\) According to Safir (1986:678), parasitic gaps offer additional proof for the existence of an operator in the COMP position in non-*wh*-relatives; see (i).

(i) every man [wh, [John saw e]] [without meeting pg]]

Notice that Borsley (1997) uses the same argument to show that the gap is a DP; cf. (11) above.

\(^5\) See further Chomsky (1977) for extensive discussion, or Van Riemsdijk & Williams (1986:93-101) for an overview.
Cross-linguistically, many relative pronouns have a \(w(h)\)-format, e.g. \textit{which/who} in English or \textit{wat/(de)welke} in Dutch. Many relative pronouns morphologically equal – or are derived from – interrogative pronouns. A second class of relative pronouns resembles demonstrative pronouns, e.g. Dutch \textit{die, dat} or German \textit{der, die, das}. Relative constructions containing one of these pronouns also meet the criteria for \(wh\)-movement. Therefore, relative pronouns in \(d\)-format also possess a \(w\)-feature. Somewhat superfluously, a further indication for this is, for instance, that Dutch \textit{die} may be interchanged with (old-fashioned) \textit{(de)welke} (with an optional \(d\) and a visible \(w\)) and that \textit{dat} may be replaced by (colloquial) \textit{wat}.\(^6\) Also some dialectal forms show a visible \(w\), instead of standard Dutch \(d\).

In short, we may state the following:\(^7\)

**Theorem I**

\(a\). Relative pronouns can have various formats: \(w(h)\)-, \(d\)-, empty or otherwise.

\(b\). All relative pronouns bear a \([+\text{wh}]\) feature.

**Theorem II**

\textit{All postnominal relative constructions undergo \(wh\)-movement of a relative pronoun.}

This concludes the discussion on \(wh\)-movement for the moment. I will return to it at several points of the discussion in later sections.

### 3.2. Case and agreement

Consider the Case and agreement facts in relative constructions. A relative pronoun agrees with the head noun, i.e. they bear the same \(\phi\)-features, i.e. number, gender and person. However, there is a possible difference in Case between the two, as is illustrated with the German sentence in (21).

\begin{align*}
\text{(21)} & \quad \text{Ich fürchte den Herrn der eine Pistole trägt.} \\
& \quad \text{I fear the.ACC gentleman.ACC who.NOM a gun carries} \\
& \quad \text{MASC.3SG MASC.3SG MASC.3SG} 
\end{align*}

Borsley (1997) states that this is a principal problem for the promotion theory, but in my view it is only a problem of execution. Kayne (1994) does not really address the issue. I will give a derivational analysis for (21) in terms of \textit{feature checking}.

First, let me clearly state in general what (21) shows for German, where both the agreement and the Case features can be seen overtly.

**Theorem III**

Relative pronouns – like nouns, determiners, and other sorts of pronouns – bear \(\phi\)-features (person, number, gender) and Case features.

---

\(^6\) The analysis of \textit{dewelke} is unclear, as yet: does it consist of one or two heads?

\(^7\) For the moment, I ignore the complications that relative particles and resumptives raise. These are treated in Chapter 5.
Consequently, relative pronouns that do not show a morphological reflex of these features (such as the English ones, or empty operators), still bear them in an abstract sense. This is of course similar to the common statement that all noun phrases bear abstract Case. Notice that in my terms a noun has its own features (including a Case feature). In Bianchi’s work, a noun has Case because it is governed by a determiner.

The next step is to establish the syntactic status of relative pronouns. In this book I follow Abney (1987) and others in that a determiner is the head of a nominal constituent. In other words: a nominal argument is a DP. The head D contains an article or some other determiner, and it selects an NP. I don’t think it would be very bold to assume that a relative pronoun is a D head as well. In fact, the standard theory of relative clauses would be perfectly happy to hear that relative pronouns are D heads. I will show that this is exactly the right assumption for the promotion theory, too.8

Theorem IV

a. Nominal arguments are DPs, where the determiner D selects NP.

b. A relative pronoun is a determiner.

The next subsection shows how the syntax of Case and agreement can be implemented in a derivational DP framework.

3.3. The relation between N and D

This subsection discusses the syntax of Case and agreement in DPs outside a relative context; §3.4 will continue with relative DPs, and show how Borsley’s problem is naturally solved. I will treat Case and agreement together, since they are closely related.

There is a clear interaction between D and N. There is not only a selectional relation. Normally, D and N also bear similar Case and agreement features. Certain determiners, namely articles, cannot even exist without a noun:

(22) a. The book is on the kitchen table.

b. * The is on the.

The tight connection between D and N can also be illustrated with a Swedish example. In (23b) the noun has incorporated into the determiner.

(23) a. ett hus [DP ett [NP hus]] ‘a house’

b. hus-et [DP hus-et [NP t]] ‘the house’

---

8 In Bianchi’s work this assumption is emphasized by the use of the term relative determiner to indicate a relative pronoun.
According to Delsing (1988/1993) there is overt N-to-D raising (incorporation of N into D) in Scandinavian. This is indicated in (23). The same kind of incorporation can be seen in other languages, e.g. Bulgarian, Basque or Erzya. Delsing assumes that there is a head raising parameter. If we submit this parameter to the logic of the Minimalist Program, it follows that there is abstract (covert) incorporation of N into D in other languages (e.g. Dutch or English). The difference between the Swedish and the English data may be encoded in the theory as a strong/weak distinction of a feature on D. If the feature is strong, N incorporates into D overtly and checking takes place. This gives (23b) for instance. If the feature is weak, only the formal features of N raise to D, as in English the house.

What kind of feature does this involve? This question is not difficult to answer. Determiners bear \( \phi \)-features, as can be seen overtly in many languages, for instance in German (cf. theorem III). If features on functional heads need to be checked – a common assumption – these \( \phi \)-features are exactly what is needed theoretically. The incorporation relation between N and dependent D makes sure that they agree and bear the same Case. If the Case and \( \phi \)-features were not compatible, incorporation would lead to a crash of the derivation. Or, from the opposite perspective: incorporation is only possible if there are no contradictory features. This, too, is a very common assumption, cf. Bianchi (2000a).

At this point, consider the necessary assumptions that are part of the derivational framework used. These are listed in theorem V, which is a summary of Chapter 1, section 3:

**Theorem V**

1. Derivations are strictly cyclic.
2. Formal features must be checked.
3. Checking of a feature (i.e. comparison with a similar feature) can take place
   (i) in a spec-head configuration, or
   (ii) in a head incorporation structure.
4. Features of a head X are visible in all projections of that head: X, X', XP.
5. Features are parametrically ‘strong’ or ‘weak’.
6. Strong features force overt checking, i.e. with PF-related (‘lexical’) material.
7. Weak features may be checked covertly, i.e. without PF-related material.
8. Incorporation does not tolerate contradictory features.
9. Excorporation is not possible.

---

9 This analysis is inspired by earlier work by Szabolcsi (1984), Hellan (1986), Abney (1987) and Ritter (1988).
10 For pre-Minimalist ideas along the same lines cf. Longobardi (1994).
11 In Swedish a definite determiner has a strong feature. (23a) shows that an indefinite determiner has a weak feature. If there is an adjective there may be doubling of the determiner. The intricacies of Swedish double definiteness are outside the scope of this discussion; but see Delsing (1993:Ch4).
12 This is prerequisite for theorem Vc, which describes checking between heads, and checking between a head and a specifier. Since a specifier contains a maximal projection according to the X'-theory, features must be visible in projections of a head. This is often described with the pretheoretical notion of percolation; see also Ch1 §3.2.
j. Covert movement is 'partial head movement' of formal features.$^{13}$

k. Covert movement is more economical than overt movement.

The consequence of Va and Vb is that formal features force movement of a constituent or head that contains the feature needed, if it cannot be obtained by merging from the lexicon (or another partial derivation).

Given this framework, consider the role of $\phi$-features in the analysis of a simple DP. A DP like *the house* in English has the surface structure (24), where the $\phi$-features of D force incorporation into D of N's formal features (FF), which include equal $\phi$-features. This creates a proper checking configuration. N's phonological features (PF) are stranded. Incorporation is indicated by a '+'.$^{14}$

\[
(24) \quad [\text{DP} \ [\text{D FF(N)+D}] \ [\text{NP PF(N)}]].
\]

Crucially, (24) is the only possible derivation that survives:

If nothing moves, i.e. the structure remains $[\text{DP} \ [\text{D N}]]$, there is no checking configuration and D's $\phi$-features cannot be checked. Therefore the derivation crashes (at the LF-interface).

If N moves overtly, the structure becomes $[\text{DP} \ [\text{D N+D}] \ [\text{NP t_n}]]$. All features can be checked. The derivation does not crash, but it is less economical than (24). If D's $\phi$-features were strong (as in Swedish *hus-et*), this option would survive.

If NP moves overtly, the structure becomes $[\text{DP} \ [\text{NP N}] \ [\text{D t_n}]]$. This is a checking configuration. Features can be checked, hence the derivation does not crash, but again, it is less economical than (24). Notice that in Swedish, this option competes with the former (that is, without further assumptions).

If there is movement but no checking, the derivation crashes (at the LF-interface), because of Theorem Vb.

Next, consider how the distribution of Case is justified. The consequence of the DP-theory is that it must be DP (the outer shell of a nominal argument) that is responsible for argument-external checking relations. In other words, checking of a Case feature on I, AgrO, AgrS, V or P (depending on the particular construction and theory) is performed by DP.$^{15}$

\[13\] This is the consequence of the idea that derivations are strictly cyclic (theorem Va). LF-movement would be countercyclic; cf. Ch1§3.2.

\[14\] X+Y can be analysed as $[Y \ X [Y]]$, which is 'head adjunction'.

\[15\] One might think that if D checks Case with a functional head in the clause, it is possible that it also checks $\phi$-features there. For instance, if DP is a subject, there is agreement with the verb, so number and person could be checked with I (or AgrS). The question is then if checking with N is still necessary. This is certainly so, e.g. because there is also a gender feature. As far as I know, verbs never show gender agreement, so D is forced to enter into a checking relation with N. In the remainder of the text I will not discuss a possible decomposition of $\phi$-features and simply state that D needs to check $\phi$-features with N.
In a nominal phrase like *the house* there are at least four relevant features: $\phi$ on N, $\phi$ on D, Case on N, Case on D. Externally, there is a Case feature on, say, AgrO. The accusative Case on AgrO must be checked, hence DP moves to SpecAgrOP. If D is not accusative, the derivation would crash, so consider the case where DP is accusative. D's $\phi$-features must be checked by N. This is described above. If these features do not match, the derivation crashes, hence suppose $\phi(D)$ equals $\phi(N)$. Regarding Case, there are in principle two possibilities: N is also accusative, or N has the 'wrong' Case, say nominative.

First suppose that N is accusative. This leads to (24) for English, or (23b) for Swedish. N is attracted by D overtly or covertly because D's $\phi$-features must be checked. Since there are no contradictory features, N may incorporate into D. Hence D's Case feature is also checked.

Second, suppose that N is nominative. If so, N cannot be incorporated into D, because there is a contradictory Case feature (cf. theorem Vh). Nevertheless, D's $\phi$-features must be checked in order to prevent a crash. Thus a less economical derivation comes into consideration: movement of NP to SpecDP:

(25) $[\text{DP} [\text{NP} N_{\text{NOM,4}}] [\text{D'} D_{\text{ACC,4}} t_\phi]]$

In this spec-head configuration the $\phi$-features can be checked. (N's nominative Case is discussed below.) D's accusative feature cannot be checked, but that is no problem, since DP as a whole moves to SpecAgrOP anyway, hence accusative Case can be checked with AgrO there.17 Thus, whether D's features are weak or strong, (25) is a converging derivation, as far as D is concerned. Still, in normal circumstances (25) will crash, because N's Case feature remains unchecked.18

However, the reader will have noticed that (25) is exactly the structure proposed by Kayne for a relative DP, cf. (7) above, where N is the antecedent and D a relative pronoun. The next subsection resumes the discussion on relative clauses and shows why (25) does not crash in a relative context. In short: because of raising, N can be linked to a higher D.19

---

16 Notice that D's features should not be 'erased' after checking (*contra* Chomsky 1995), because they are still necessary, because, when the derivation proceeds, AgrO must be checked with DP. Therefore, I assume that D's checked features receive a check-mark, as argued also in Ch1§3.2.

17 Recall from Ch1§3.2 that a feature *can* be checked in a spec-head configuration. This implies that not all features *need* to be checked. Complete spec-head agreement ("SHAGR") is not a desirable theoretical concept, because it blurs the difference between incorporation (head movement) and spec-head relations (XP-movement). This is not difficult to prove. If spec-head agreement were to involve all features, it would in general not be possible to generate an XP in SpecVP, or to move it there, because the categorial and other features (almost) never match. For instance, if a DP [+N,-V] is generated as the external argument of a verb in SpecVP [-N,+V] the derivation crashes before it even starts, so to speak.

18 Theorem Vb states that formal features must be checked. In the Minimalist Program this is often interpreted as: only formal features of functional heads must be checked. I reject this additional assumption, see also Ch1§3.2.

19 Notice that the view presented here differs in some respects from De Vries (1996).
3.4. A detailed derivation of postnominal restrictive relatives

Given the above premises, the derivation of a relative clause like (21) is straightforward. This is illustrated in (26), which is a merge-and-move schema, starting with the embedded clause. A detailed explanation is given immediately below. I will discuss the derivation of this German sentence first and consider other possibilities right after that. Notice that the derivation in English is exactly the same.

(26) Ich fürchte den Herrn der eine Pistole trägt.

'I fear the gentleman who carries a gun.'

\[ \text{DP-rel der [NP Herrn]} \rightarrow \]
\[ \text{DP-rel [NP Herrn] der [NP Herrn]} \rightarrow \]
\[ \text{VP [DP-rel Herrn der [NP Herrn]} \left[ V \text{ eine Pistole trägt}] \rightarrow \]
\[ \text{IP [DP-rel Herrn der [NP Herrn]} \left[ I \left[ V \text{ eine Pistole trägt}] \rightarrow \]
\[ \text{CP [DP-rel Herrn der [NP Herrn]} \left[ C \left[ I \left[ V \text{ eine Pistole trägt}] \rightarrow \]
\[ \text{IP den [CP [DP-rel Herrn der [NP Herrn]} \left[ C \left[ I \left[ V \text{ eine Pistole trägt}] \rightarrow \]
\[ \text{DP FFh+den [CP [DP-rel Herrn der [NP Herrn]} \left[ C \left[ I \left[ V \text{ eine Pistole trägt}] \rightarrow \]
\[ \text{CP Ich fürchte [DP FFh+den [CP Herrn der eine Pistole trägt]} \rightarrow \]

The (future) head noun Herrn originates in the subordinate clause, viz. as the complement of the relative pronoun der; see (26a). This is one of the basic assumptions of the raising/promotion analysis. It reflects the fact that the head noun plays a semantic role in the relative clause. D_{rel} and N will be ultimately disconnected in a way, so they do not necessarily match in every respect (read: their Cases may be different). Nevertheless, the categorical selection of NP by D is as usual; moreover they must agree in \( \phi \)-features.

The \( \phi \)-feature agreement between D_{rel} and N needs to be checked. This licences movement. In this context there cannot be N-to-D raising, since N and D_{rel} have a contradictory Case feature (i.e. D_{rel} is nominative, N accusative). Therefore incorporation is not allowed. Still, D_{rel}’s \( \phi \)-features need to be checked. Therefore NP is attracted to SpecDP_{el} in accordance with theorem Vc(i) above; see (26b). This is a normal checking configuration and the \( \phi \)-feature agreement is settled. The contradictory Case features cannot be checked at this point. Structure (26b) is like (25). If nothing else happens, the derivation would crash because of unchecked Case features. Notice that although D’s \( \phi \)-feature is weak in German, the system forces overt movement of NP in this context. As explained, incorporation — i.e. (abstract) head movement — is not an option here. Hence a less economical derivation must be chosen, because the more economical one would crash; cf. Ch I §3.2.

Next, DP_{rel} as a whole is selected as the subject of the predicate eine Pistole trägt in (26c). Thus the requirement that an argument position must be occupied by a DP is fulfilled.

---

20 Here, a spec-head-comp order is assumed in the functional domain. The argument abstracts away from the VO/OV discussion in Dutch and German. See further section 3.6.
When I (or AgrSP if one likes) is merged with VP in (26d), it attracts the subject DP_{rel} in order to check nominative Case (probably along with the EPP – if that is a feature – and other features, e.g. person/number agreement). Obviously, the Case associated with N (which is accusative) does not enter in any checking relation at this point.

In (26e) the CP-level is added. The relative pronoun possesses a wh-feature; therefore DP_{rel} moves to SpecCP and the wh-features are checked in spec-head configuration.

Then, in accordance with the D-complement hypothesis, the whole CP is merged as the complement of D (den); see (26f).

At this point it becomes clear why a structure like (25) is allowed in a relative clause. The derivation does not crash, because N can be associated with a higher determiner. Since the head noun is in the specifier of DP_{rel} in the specifier of CP, there is no barrier between D (den) and N (Herrn). D’s φ-features attract N. In German these features are weak, so the movement is covert (because that is more economical than overt movement). This is indicated in (26g) by moving the formal features of N to D. Hence the agreement between D and N is checked in an incorporation structure. Incorporation is only possible if all features match, therefore N and D must also bear the same Case – accusative in this example. Hence the Case features are checked, too.

Finally, the whole DP is inserted into the matrix clause in (26h). DP is accusative, since its complex head FF_h+D is accusative. Hence the matrix clause AgrOP (not indicated in (26)) can check its accusative Case feature with DP.

Although D and N originate separately, they end up together. They agree and bear the same Case, which may differ from the Case of the relative pronoun.

Next, some relevant other possible derivations will be considered. First, take a Swedish sentence like (27).

(27) Jag talade med mann-en vilken känner dig.
    I spoke with man-the who knows you
    ‘I spoke with the man who knows you.’

Up to (26g) the derivation exactly parallels the one in German or English. The final two steps are sketched in (28).

(28) a.-f. ... compare (26a-f)...
    g. [DP mann_{h+en} [CP [DP_{rel} t_h vilken t_h], C [DP t’, I [VP t, känner dig]]]] →
    h. [CP Jag talade med [DP mann_{h+en} [CP vilken känner dig]]] →

See also Kayne (1994), Barbiers (1995), De Vries (1996), and Bianchi (1999/2000a). Notice that there is no intermediate head.
Since the outer D has strong $\phi$-features, the head noun *mann* incorporates overtly into this determiner *en*.22

Second, consider the possibility that all relevant Cases happen to be equal, as in *the NOM man NOM who NOM saw you, left*. The relative DP is [[D-rel who] [NP [N man]]] initially. Then, there are three options:

(i) N incorporates into D-rel overtly.
(ii) N incorporates into D-rel covertly, i.e. the formal features of N move to D-rel.
(iii) NP moves to SpecDP-rel.

Each of the three options is a valid step at this point of the derivation, because it leads to a checking configuration; hence the $\phi$-features and nominative Case are checked. However, in Swedish (ii) would not be allowed because D’s features are strong. In English, (ii) will eventually be preferred over (i), because it is more economical. But the distinction between (i) and (ii) on the one hand and (iii) on the other is more interesting. When the derivation proceeds, DP-rel moves to SpecCP. D-matr(\text{ix}) selects CP-rel. At this point the relation between N and D-matr is crucial. D-matr needs to check its $\phi$-features with N, so the latter is attracted. Therefore N incorporates into D and the features are checked. In Swedish this is overt, in English covert. If initially step (iii) was taken, this is no problem: the whole procedure is similar to the one described in (26). However, if initially (i) or (ii) was performed, N or FF(N) has to excorporate from D-rel before it can incorporate into D-matr. Many linguists assume that excorporation is not a possible step (cf. theorem Vi). If so, the derivation built from (i) or (ii) eventually crashes, because N is locked in D-rel, so D-matr’s $\phi$-features remain unchecked. Therefore the derivation starting with (iii) – although less economical than (ii) – is probably the only survivor in a relative context.

Third, consider the hypothetical possibility that three different Cases are selected: ...D-matr\text{DAT} N\text{NOM} D-rel\text{ACC} ..., which is wrong. Since the Cases do not match, incorporation is impossible. Therefore initially NP moves to SpecDP-rel and the $\phi$-features are checked. DP-rel checks Case in the clausal domain, in this example with AgrO. DP-rel moves to SpecCP. D-matr selects CP. Then NP moves to SpecDP-matr, so that D-matr’s $\phi$-features can be checked. (Again, incorporation is impossible because dative and nominative do not match.) DP-matr will check dative in the matrix clause. The relevant configuration is now (29), (where NP has been raised from within CP):

(29) [[D-matr [NP N\text{NOM}] [D\text{DAT} CP\text{rel}]]

This structure is comparable with (25) above. I have shown that it will survive only if N can be associated with a higher D. Since there is no such D available in the

---

22 As noticed before, the option that NP moves to SpecDP instead of overt N-into-D incorporation cannot be excluded a priori. It establishes the same word order and an equivalent feature checking configuration. Nevertheless, an argument for head movement could be the fact that *manner* is phonologically one word.
matrix, the derivation will crash – as desired – because N’s nominative Case feature remains unchecked.

Finally, I must mention that there is one possible derivation which gives a wrong result, but which I do not know how to exclude without further assumptions. Suppose N, Dmat and Drel have equal φ-features; N and Drel are nominative, and Dmat is accusative (or the other way round): e.g. *den Herr der...theACC gentlemanNOM whoNOM...*. Initially NP moves to SpecDPrel; the φ-features and nominative Case (!) are checked. DPrel moves to SpecCP; CP is selected by Dmat. Overt or covert incorporation is impossible, because the Cases do not match, hence NP moves to SpecDPmat. D’s φ-features are checked. The Case of DPmat is checked with AgrO in the main clause. Thus all features are checked and the derivation survives. The problem compared to (29) is that here N’s Case feature has already been checked in the relative clause, so it cannot cause a crash. A potential solution is to assume that if NP is raised into a new clausal domain, its features must be re-licenced; but I will leave this issue open for further discussion.

3.5. ‘That’-relatives

What happens if there is no relative pronoun, as in (30b)?

(30) a. I fear the gentleman who carries a gun.
    b. I fear the gentleman that carries a gun.

In my view these small variations in the COMP area are only surface effects. I prefer to treat the data in a uniform way, following in fact Chomsky (1977). The sole difference between (30a) and (30b) is that (30a) has an overt relative pronoun, whereas (30b) has an overt complementizer. Chomsky argues that (30b) has an empty operator which is the equivalent of a relative pronoun. The Doubly Filled COMP Filter makes sure that the relative pronoun and the complementizer cannot be overt both in standard English. See the next chapter for more discussion.

In the promotion theory the empty operator is represented as the determiner Drel. It is phonetically empty, but it does have all the formal features of a relative pronoun, i.e. at least Case and φ-features and a wh-feature. Therefore the analysis for (30b) exactly parallels (30a) and (26).

Borsley (1997) objects that if there is an empty Drel, there is a danger of it being filled with an article, as in (31). This would lead to a doubled article, superficially.

(31) *the [CP [DP-rel the gentleman], that I saw t,]  

However, (31) is excluded for obvious reasons. An article is never [+wh], hence raising would not be possible at all. Therefore the external article the remains without a noun, and its φ-features cannot be checked. So the derivation crashes.

Borsley also objects to what seems to be the reverse of the same problem: if there can be an empty D in (30b), why is an empty determiner in a non-relative definite DP impossible, e.g. *(the) man? The answer is, I think, that the question is
wrong. The phonetically empty D in (30b) is not an arbitrary D, it is an empty relative pronoun $D_{e}$. In Dutch, there are no empty relative pronouns. Clearly, this is only a lexical difference. English happens not to have empty 3SG articles. Other languages do; Latin for instance.

3.6. Word order variation

In Chapter 3 I have stressed that many different word orders with respect to relative constructions have to be accounted for. Relevant is the basic word order of a language and the linear order of the external determiner, the head noun and the relative clause (which itself may contain a relative pronoun, et cetera). I have shown that all possible permutations are actually attested. For postnominal relatives, the patterns are repeated in (32) and (33). See also Appendix II, table 16.

(32) SVO languages
  a. D N R C e.g. in English
  b. N D R C e.g. in Swedish
  c. N R C D e.g. in Indonesian

(33) SOV languages
  a. D N R C e.g. in Hindi
  b. N D R C e.g. in Oromo
  c. N R C D e.g. in Lakota

The patterns in (32a) and (32b) have been treated extensively in the previous sections, the other ones remain to be explained.

In the previous chapter I have ‘calculated’ the movements necessary to derive these linear orders within several theories on syntactic structure. For the promotion theory the results are summarized in table 1. (See Ch3§3.2 for definitions, etc.)

Table 1. Movements in promotion theories required to derive word order variations in postnominal relative constructions.

<table>
<thead>
<tr>
<th>Subtheory</th>
<th>Antisymmetry, rigid left, func left-local uniform branching</th>
<th>Uniform branching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VO and OV spec D [cp, NP RC]</td>
<td>[cp, RC, NP] D spec</td>
</tr>
<tr>
<td>Linear order</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D N R C</td>
<td>√</td>
<td>$H_{a}^{d} + rM_{dp}$</td>
</tr>
<tr>
<td>N D R C</td>
<td></td>
<td>$rM_{dp}$</td>
</tr>
<tr>
<td>N R C D</td>
<td></td>
<td>$A_{ep}$</td>
</tr>
</tbody>
</table>

With respect to the promotion theory of relatives, we can see that an antisymmetric phrase structure, but also ‘rigid left’ and ‘fl-lub’, do not distinguish VO languages from OV languages in any relevant sense. The surface linear order D N R C (the most common type) is derived in the way argued for in the previous sections (cf. (26) above); the order N D R C requires additional overt head movement of N to D,
as explained (cf. (28) above). The pattern N RC D can only be shown if CP with all its content is moved to SpecDP. The resulting structures are shown in some detail in (34), where the determiner and the head noun are printed in bold, and the relative clause (RC) is underlined.

(34) a. \[ \text{DP} \left[ \text{DF(N)+D} \right] \left[ \text{CP} \left[ \text{DP-rel [NP N]} \left[ \text{D-rel} \right] \left[ \text{t_{DP}} \right] \text{[C]} \left[ \text{[IP ... t_{DP} ...]} \right] \right] \right] \]

b. \[ \text{DP} \left[ \text{D N+D} \right] \left[ \text{CP} \left[ \text{DP-rel [NP t_{n}]} \left[ \text{D-rel} \right] \left[ \text{t_{DP}} \right] \text{[C]} \left[ \text{[IP ... t_{DP} ...]} \right] \right] \right] \]

c. \[ \text{DP} \left[ \text{CP [DP-rel [NP N]} \left[ \text{D-rel} \right] \left[ \text{t_{DP}} \right] \text{[C]} \left[ \text{[IP ... t_{DP} ...]} \right] \right] \right] \left[ \text{CP t_{CP}} \right] \]

The N RC D pattern can be represented as in (34c). Consider how the derivation might proceed. As discussed before, NP moves to SpecDP, and the \( \phi \)-features are checked. DP moves to SpecCP and \( \text{wh} \) is checked. CP is selected by D. D needs to check \( \phi \) with N (and N needs to check \( \phi \) and possibly Case with D), so there must be movement of some kind. At this point the derivation deviates from (34a/b). Instead of movement of NP to SpecDP (or incorporation of N into D), a large constituent in which N is contained moves to SpecDP, namely CP. Why is this possible? I don’t think CP has a particular reason to be in SpecDP. A solution may be sought in the concept of pied piping. According to Koster (2000a) essential differences between languages are differences in pied piping. Although Koster’s ideas are much more far-reaching than can be discussed here (he suggests that the whole overt/covert distinction can be captured by differences in pied piping), I propose to allow for at least some parametrization in pied piping.

Well-known differences with respect to pied piping are examples with prepositions; see (35).

(35) a. \[ \text{PP In which city} \text{ does he live t_{PP} ?} \]

b. \[ \text{DP Which city} \text{ does he live [in t_{CP}] ?} \]

Some languages prefer (35a), some (35b), and in some pied piping is optional. Of course pied piping is influenced by opacity effects. If PP is a barrier, (35b) is not an option.

Thus suppose – more or less in the spirit of Koster (2000a) – that a particular language prefers pied piping instead of N(P) movement. Then the whole CP in which NP is contained may be moved (similarly to the fact that PP is moved instead of DP in (35)). This is a pretheoretical statement, of course. Therefore consider what it means in terms of feature checking. Feature checking in a spec-head configuration is checking between a maximal projection XP in the specifier position of a head Y, and Y itself. So if XP is a pied piped constituent, the relevant features of an embedded head must have percolated to the head X (and consequently to XP, as noted before). For instance, in (35) the wh-feature of which has percolated up to P(P). The PP moves to SpecCP and checks \text{wh}.

In (34c) pied piping means that N’s \( \phi \)-features and Case percolate up to C(P). So CP moves to SpecDP and these features can be checked.

It is also clear why (36) is excluded. The derivation of (36) would force pied piping of DP_{rel} instead of CP:
The Promotion Theory of Relative Constructions 129

But that would mean that N's features percolate up to DP_{rel}, which already has the same type of features of its own. (Moreover, these may be contradictory.) I have argued before that features are not 'erased'. The idea that N's features can overwrite D_{rel}'s features (or duplicate if they happen to match) seems highly implausible to me. Hence suppose the following:

Theorem VI
Pied piping can be the result of feature percolation to a higher head (or projection) which itself does not bear this kind of features.\(^{23}\)

Thus typical nominal features like Case or \(\phi\) may percolate up to C(P), or \(\mathit{wh}\) may percolate to P(P), since these categories do not inherently bear this kind of features themselves.

These suggestions relate a particular word order regarding relative clause constructions (34c) to pied piping. A thorough examination of pied piping phenomena in the relevant languages is needed to further substantiate this kind of claims. This, however, is far beyond the scope of this book.

For now, this concludes the discussion of three versions of the promotion theory (antisymmetry, rigid left and fl-lub) with respect to the derivation of word order variation in postnominal relatives. As can be seen in table 1, the same conclusions can be drawn for the uniform branching version in VO languages. Thus what remains to be discussed is uniform branching with respect to postnominal relatives in OV languages. The surface representations for the three word orders, based on table 1, are given in (37). Recall that heads and specifiers are on the right here.

\[
(37) \quad \text{Can these representations be derived in a plausible way? Consider (37a) first. Within CP the usual movements and checkings are performed. After selection of CP by D, N is incorporated into D. It is merged to the right of D, which is plausible since movement is also to the right due to right-hand specifiers. So Case and \(\phi\)-features can be checked between N and D. After this, there is remnant movement of CP to SpecDP. I can think of no plausible trigger for this. Moreover, if there is a relative pronoun or particle it would be on the right (i.e. clause-final). However, Appendix II, table 16 or 24, shows that this pattern is not attested, whereas there are several}
\]
counterexamples where $D_{rel}$ or $C$ are clause-initial, e.g. in the Indo-Aryan languages Hindi, Bengali and Marathi, and in the Germanic languages Dutch, German and Frisian. This shows that it is not a viable strategy to have right-hand specifiers, at least in these languages. This casts serious doubt on the uniform branching hypothesis as defined.

Nevertheless, consider (37b). This, too, is not a plausible representation. The position of relative elements is structure-initial, which is impossible. Therefore another strategy must be found. Suppose that $N$ moves to $D$ and left-incorporates. If so, there can be CP remnant movement to SpecDP instead of IP remnant movement. This gives us (37a) again, except that D+N is replaced by N+D. Hence the potential problems are the same. There is no trigger for the CP remnant movement involved, and the position of relative elements would be clause-final. The clause-initial relative particles in Farsi, Urhobo and Oromo show that this is wrong.

The pattern $N$ $RC$ $D$ in (37c) can only be derived by moving NP leftwards and leaving CP in situ as the left-hand complement of $D$. There is no regular landing position for NP, so it can only be left-adjoined to CP – an unmotivated movement. There is still no checking relation between $N$ and $D$, so there must be additional (covert) formal feature movement of $N$ to $D$. This is not indicated in (37c). Again, the position of relative elements is clause-final. There are neither examples nor counterexamples in the data set.

Thus, word order variations in postnominal relatives can be derived in an antisymmetric, rigid left or fl-left promotion theory, but not in a uniform branching theory (that is, for OV languages). The latter uses unmotivated movements and, even worse, predicts clause-final relative particles, which is at variance with the data. Theories with left-hand specifiers and left-hand functional heads do not face this problem. The (relatively rare) pattern $N$ $RC$ $D$ involves a special instance of pied piping. This seems to be an interesting phenomenon which requires more study in general.

3.7. Conclusion

In short, the various movements needed for the promotion theory can be derived straightforwardly within a framework based on feature checking. All Case and $\phi$-feature agreement facts are accounted for in a coherent way. No additional features or projections are needed. Moreover, all restrictive postnominal relatives are analysed alike, regardless of the particular setting of the COMP area (which itself is discussed in more detail in Chapter 5). It seems that the cross-linguistic word order variations regarding head noun, external determiner and relative clause can be explained only if functional heads and specifiers are on the left. I will return to this issue.
4. Prenominal relatives

Prenominal relative constructions are rarer than postnominal ones. Nevertheless they occur in different language families around the world; cf. Appendix II, figure 1 and table 5.

In principle, the analysis of postnominal restrictive relatives presented in the previous section can be extended to prenominal ones straightforwardly. The major difference with postnominal relatives is that the order of the head noun and the relative clause is switched. There are roughly two ways to derive this. Either there is a leftward branching comp-head-spec scheme, or there is an additional movement of the relative clause to a position left of the head noun. This is sketched in (38), where the head noun is in bold face and the relative clause is underlined.

\[(38) \text{ a. \left[ DP \left[ \text{CP .. \ldots t_{DP .. \ldots NP}} \text{ (D)} \right] \right] (\text{comp-head-spec} \text{ base})} \]
\[\text{ b. \left[ DP \left[ \text{.. \ldots t_{DP \ldots \ldots}} \text{ (D) [CP NP t]} \right] \right] (\text{spec-head-comp} \text{ base})} \]

I will discuss (38) in detail below.

Three additional properties of prenominal relatives are stated in (39), repeated from Ch2§5; see also Appendix 2, tables 5/8/10/11.

\[(39) \text{ a. Prenominal relatives do not have relative pronouns.} \]
\[\text{ b. Prenominal relatives do not have clause-initial relative particles.} \]
\[\text{ c. If there is a clause-final relative particle, it does not equal the regular complementizer.} \]

According to Kayne (1994:92-95) this is no coincidence. In his theory, which is like (38b), the prenominal relative is a raised IP. Since relative pronouns and complementizers are in SpecCP and C, there can be none in a prenominal relative IP. Thus at first sight it seems favourable to derive prenominal relatives from postnominals. Unfortunately, if we look at the details of the derivations, this is much less clear.

The position of the determiner may be important. Many languages do not have a regularly overt determiner, but several do. Therefore consider the possible word order variation. Prenominal relatives predominantly occur in OV languages, as one would expect. In previous chapters I have already mentioned that all permutations of the external determiner, the head noun and the relative clause are attested, cf. (40). See also Appendix 2, table 17.

\[(40) \text{ SOV languages} \]
\[\text{ a. DRCN e.g. in Tigré} \]
\[\text{ b. RCDN e.g. in Korean} \]
\[\text{ c. RCND e.g. in Basque} \]

---

24 Considering that Mandarin Chinese has prenominal relatives, we cannot conclude that there are few speakers of this variant.
Chapter 4

There are some examples of prenominal relatives in VO languages (cf. Appendix 2, table 23). Unfortunately, in the language sample I have compiled, there is little information on the position of determiners in these languages. There is no example of a language that regularly uses a definite determiner. Nevertheless, there are examples with a quantifier or (demonstrative) determiner in Mandarin Chinese; these can be in an initial or middle position.25

(41) SVO languages

(a) D RC N (e.g. in Chinese) e.g. Palauan, Finnish or Chinese
(b) RC D N (e.g. in Chinese)
(c) RC N D (not in the sample)

Recall from the previous section that D can take any position in postnominal relative constructions, whether in OV or VO languages. I am convinced that if more data becomes available, there will be clear(er) examples of all three patterns in (41), too. Therefore I will treat these on equal terms with (40), which implies that a theory on relative clause structures should be able to derive them.

Table 2 summarizes the (additional) movements necessary to derive the patterns above within a promotion theory of relative clauses, as discussed in Ch3§3.2. Again, antisymmetry, rigid left and fl-lub do not distinguish a different ‘basic order’ for OV and VO languages in this respect, hence (40) and (41) are treated as one group, which obliterates the lack of data concerning (41). Uniform branching on the other hand does make a distinction.

Table 2. Movements in promotion theories required to derive word order variations in prenominal relative constructions.

<table>
<thead>
<tr>
<th>subtheory →</th>
<th>antisymmetry</th>
<th>rigid left, func left-l.u.b.</th>
<th>uniform branching</th>
</tr>
</thead>
<tbody>
<tr>
<td>word order</td>
<td>VO and OV</td>
<td>spec D [CP NP RC]</td>
<td>[RC NP CP D] D spec</td>
</tr>
<tr>
<td>‘basic’ structure →</td>
<td>VO</td>
<td>[CP NP RC]</td>
<td></td>
</tr>
<tr>
<td>linear order ↓</td>
<td>VO</td>
<td>[CP NP RC]</td>
<td></td>
</tr>
<tr>
<td>D RC N</td>
<td>??</td>
<td>rA_{dp} \uparrow</td>
<td></td>
</tr>
<tr>
<td>RC D N</td>
<td>rM_{dp} \uparrow</td>
<td>M_{\uparrow}^+ \uparrow</td>
<td></td>
</tr>
<tr>
<td>RC N D</td>
<td>H_n^+ \uparrow + rM_{\text{spec}}^+ \uparrow</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Consider first the detailed derivation of prenominal relatives in OV languages within the uniform branching theory. The structures are in (42). Recall that they are left-branching, i.e. comp-head-spec.

(42) a. [DP \text{t}_{cp} D [\text{CP} \ldots \text{t}_{dp-rel} \ldots (C) [\text{DP-rel} t_{np} (D_{rel}) [NP N]_{dp-rel} t_{cp} [D D+FF(N)]]]

   b. [DP \text{t}_{cp} D [\text{CP} \ldots \text{t}_{dp-rel} \ldots (C) [\text{DP-rel} t_{np} (D_{rel}) [NP t_{i}]_{dp-rel} t_{cp} [D D+FF(N)]]]

   c. [DP \text{t}_{cp} D [\text{CP} \ldots \text{t}_{dp-rel} \ldots (C) [\text{DP-rel} t_{np} (D_{rel}) [NP N]_{dp-rel} t_{cp} [D D+FF(N)]]]

The structure in (42c) is the exact mirror of the English postnominal one in (34a). In (42b) there is overt head movement – incorporation of N into D – as in Swedish (34b). In (42c) there is pied piping of the whole CP to SpecDP as in Indonesian (34c). I will not repeat the discussion concerning feature checking here. Notice that, given the properties in (39), C and D_rel must be abstract. The derivations presumed in (42) are attractive in a way, because they mirror the ones for postnominal relatives and can be explained in the same way. However, the properties in (39) remain unexplained.

The derivations of prenominal relatives in VO languages (with spec-head-comp bases) are more difficult. The structures are given in (43). Notice that within the rigid-left, fl-lub and (except for (43a)) antisymmetry theories all prenominal relatives (in VO and OV languages) must be derived like this.

(43) a. \([\text{DP} [\text{DFF} (N)+D] [\text{CP} [\text{IP} \ldots \text{t}_{\text{dPred}} \ldots]] [\text{CP} [\text{DPPrel} [\text{NP} N] (\text{D}_{\text{rel}}) \text{tp}_{\text{dPred}} (C) \text{tp}]]]\]

b. \([\text{DP} [\text{IP} \ldots \text{t}_{\text{dPred}} \ldots]] [\text{DFF} (N)+D] [\text{CP} [\text{DPPrel} [\text{NP} N] (\text{D}_{\text{rel}}) \text{tp}_{\text{dPred}} (C) \text{tp}]]\]

c. \([\text{DP} [\text{IP} \ldots \text{t}_{\text{dPred}} \ldots] [\text{NP} N+D] [\text{CP} [\text{DPPrel} [\text{NP} t_p] (\text{D}_{\text{rel}}) \text{tp}_{\text{dPred}} (C) \text{tp}]]\]

c'. \([\text{DP} [\text{CP} [\text{DPPrel} [\text{NP} t_p] (\text{D}_{\text{rel}}) \text{tp}_{\text{dPred}} (C) \ldots \text{t}_{\text{dPred}} \ldots] [\text{CP} [\text{D} N+D] \text{tp}]]\]

First consider the derivation in (43b). It equals the one in English (cf. (34a) above), except for one final additional step: remnant movement of the relative clause to SpecDP. This derivation is the detailed variant of Kayne’s proposal for prenominal relatives. I would like to make three remarks here. First, notice that it only describes one out of three possible word order variants. Second, the final remnant IP movement seems unmotivated. Third, the (abstract) relative pronoun and complementizer are stranded – because i) \(D_{\text{rel}}\) and IP do not form a constituent, and ii) an \(X\)'-level (here \([C \ C \ IP]\)) cannot be moved – but the linear order RC D N D_{rel}/C is never attested, hence it must be explained why \(D_{\text{rel}}\) and C are never overt in this context. In some footnotes, Kayne acknowledges this third problem. As for the complementizer, he suggests that there may be a that-trace effect. However, the that-trace effect is by no means universal, whereas the prohibition *RC D N C does seem to be so. As for the relative pronoun, Kayne suggests that it cannot be interpreted if it is stranded. If so, it must be possible to prove that there cannot be reconstruction in this context. However, if there is no reconstruction, the relative clause would be interpreted as appositive, which is not intended. For, the external determiner must take scope over the relative clause in a restrictive relative (see also Chapter 6). This is not the case in the surface structure of (43b). Moreover, it seems to me that an abstract relative pronoun must also be interpreted.

For the moment, I will put these problems aside and proceed with (43c), which is a variant of (43b). To switch the linear order between D and N, N must overtly incorporate into D. This process has been discussed repeatedly above. Next, we may move IP to SpecDP in (43c), as in (43b). Another possibility is to move the whole remnant CP to SpecDP, as in (43c'). This latter option looks familiar. Movement of CP to SpecDP – cf. (34c) – has been described in terms of pied piping. N’s formal

---

26 Notice that in my theory there is no reconstruction, but rather cyclic feeding of the LF-component; cf. Ch1§3.2.
features must be checked with D. If they percolate up to the CP-level, CP moves to SpecDP. If both CP moves and N incorporates into D (as is the case here), it looks as if something is done twice. Double marking does sometimes occur in the grammar, but I do not wish to treat (34c') as heavily marked.

There is an alternative view, however. It is not one single feature that drives this process: a bundle of features (Case and $\phi$-features) is involved. Suppose that one (group) of these percolates up and the other(s) remain where they belong – on N. This forces two movements: N head movement to D in order to check the remaining feature(s), and CP remnant movement to SpecDP in order to check the percolated feature(s) with D in spec-head configuration. So the formal features are not treated as a solid group. This is in contradiction with usual conventions, but I think nothing in the system actually prevents it.27

If so, there is a possible explanation for the movements in (43c'). On the contrary, there is no explanation for IP remnant movement as in (34c), because IP does not contain NP. It seems particularly implausible to me to assume that there is upward percolation to the main projection line first, and then downward percolation to IP. So assume that features cannot percolate down in general.

**Theorem VII**
*Feature percolation is unidirectional. Since it starts from a head, there is only upward percolation. This causes pied piping.*

There may be other advantages of (43c'). Recall that a specifier of XP $c$-commands the head X. According to Kayne's definitions, a specifier of a specifier also $c$-commands this head. If so, in (43c') DP$_{rel}$ c-commands [D N+D]. Hence, if it is overt, (the head of) DP$_{rel}$ is a (relative) pronoun referring to N, so Principle C of the Binding Theory is violated.28 That explains (39a): prenominal relatives do not have relative pronouns. Unfortunately I don’t see a straightforward explanation of (39b): prenominal relatives do not have clause-initial relative particles. As for (39c) – if there is a clause-final relative particle, it does not equal the regular complementizer – it may simply be noted that C is not clause-final. Hence a clause final particle cannot equal a regular complementizer.

As explained, (43b) cannot be justified like this. A way out may be the following: suppose that there can be incorporation to the right. If so, we can generalize over RC N D and RC D N. The derivations are exactly the same, except for one thing: in the latter case N incorporates to the right of D. This is shown in (44), which replaces (43b).

\[
\text{[DP [CP [DP$_{rel}$ [NP $t_{D}$] (D$_{rel}$) $t_{D}$]$_{rel}$ (C) ... $t_{rel}$ ...]$_{rel}$]$_{rel}$ [D+N] $t_{NP}$]}
\]

27 This idea does not affect the results of the previous section. The relevant case is (34c), where there is pied piping of CP to SpecDP. If there were a split process, whereby N moves to D and the relative clause moves to SpecDP, the outcome would be a prenominal relative as in (43c'). Hence this parameter setting is not available. See also section 7.

28 See De Vries (1998a) for a discussion of Principle C in a derivational grammar. Until now, Principle C has always been used with respect to XP positions. Here, it is tentatively proposed that it may also apply to argument heads.
This structure departs from Kayne’s original proposal, but it has the advantage that there is a possible explanation for the movements, and the properties in (39).

Finally, consider the D RC N order, as in (43a), repeated in (45). Here the difficulty is that the relative clause splits D and N apart.

\[(45) \ [D P [D F F(N)+D] [C P \{IP \ldots t_{IP-rel} \ldots\} [C P [D P-rel [NP N] (D_{rel}) t_{IP} ]_{IP-rel} (C) t_{IP}]]]\]

So the relative IP must move to a position between D and N. Unfortunately, this position does not exist. Therefore adjunction is necessary (which is impossible in a strict antisymmetric structure), unless an intermediate projection FP is assumed, as shown in (46).

\[(46) \ [D P F F(N)+D [F P \{IP \ldots t_{IP-rel} \ldots\} F [C P [D P-rel [NP N] (D_{rel}) t_{IP}] (C) t_{IP}]]]\]

In either case it remains to be explained i) what the motivation of IP-movement is; ii) why the stranded D_{rel} and C must be empty; and iii) how the relation between D and N can be established, i.e. how the formal features of N can move to D if there is intermediate material. I don’t see how to answer these questions.

The nature of FP, if necessary, raises additional issues. In principle, it could belong to the DP domain or to the CP domain. That is, F could be a Q head or a split-CP head such as Topic or Focus. If one chooses the first option, the D-complement theory must be reconsidered. The second option may be in contradiction with the sentence-initial nature of relative pronouns and complementizers in general. Hence (46) is not very plausible.

I conclude that there is an attractive set of derivations available for prenominal relatives in OV languages within the uniform branching analysis. These exactly mirror the derivations for postnominal relatives in VO languages. With some additional assumptions, prenominal relatives in VO languages can also be derived. There is one remaining problem: prenominal relatives in VO languages in which the RC splits the determiner and noun apart.

The antisymmetric, rigid left and fl-lub theories treat prenominal relatives in OV and VO languages on a par. This has the consequence that there cannot be mirror-derivations: all derivations are like those described for VO languages in the uniform-branching theory. This puts more weight on the remaining problem for D RC N structures.

5. Circumnominal relatives

This section discusses the syntax of circumnominal relatives. It is divided into four parts. Section 5.1 introduces the topic and summarizes the properties of circumnominal relatives; 5.2 sketches the history of the analysis; 5.3 is on the derivation of circumnominal relatives within the present framework; and 5.4 briefly discusses some additional issues.
5.1. Introduction and properties

Culy (1990:27) states: "A restrictive internally headed relative clause is a nominalized sentence which modifies a nominal, overt or not, internal to the sentence."

Circumnominal relatives (often called IHRCs) have the following appearance (where the head noun and the external determiner are in bold face and the relative clause is underlined):

\[(47) \ [\text{DP} \ [\text{NP} \ldots \text{N} \ldots ] (\text{D})] \]

In short, they have the following properties:

(48) a. Circumnominal relatives are nominalized sentences, i.e. DPs.
   b. The head noun is in situ.
   c. As for relative elements:
      (i) there is no relative pronoun or marker;
      (ii) there is no resumptive pronoun;
      (iii) there is no relative complementizer (except perhaps in Dagbani);
      (iv) there can be a relative affix.
   d. As for word order:
      (i) circumnominal relatives occur in SOV, SVO (and other) languages;
      (ii) they occur in N D languages, and in D N languages that do not regularly use an overt determiner.
   e. As for the external determiner:
      (i) If D is visible, it follows the relative clause.
      (ii) The use of determiners that must be pronounced is shunned in D N languages. Hence, D can only be overt in N D languages.
   f. The internal head must be indefinite.

Regarding the syntax of circumnominal relatives, I will follow Culy's basic insights. Culy (1990:73-79) argues that a circumnominal relative is of category N'. This is because determiners, Case morphemes and other particles, if present, always follow the relative clause. This has been discussed before in Ch2§6.2. Thus the structure is like (49a), where a possible determiner is at the rightmost triple of dots. Culy (1990:68) notes that the exocentric nature of this representation can be overcome if a DP structure is used as in (49b). This is in line with the theory presented in this book.

---

29 See Ch2§5, 6.2, Appendix II and Culy (1990). In particular, concerning (48a) see Ch2§6.2 and Culy (27-29, 200-206, 264); for (48b) cf. Appendix II, table 3; for (48c) cf. Appendix II, tables 3 and 8-15, and Culy (1990:70-72); for (48d) cf. Appendix II, tables 3/21, and Culy (1990:207-229, 261); for (48e) see Appendix II, table 18, and Culy (1990:207-229, 261); for (48f) see Culy (1990:167-181).

30 Except for Yavapai, where the head noun is fronted. See below.

31 Note that the ordering of D and N is not necessarily correlated to the O/V parameter.
(49) a. \([\text{NP} \ldots [\text{NP}' \ldots \text{NP}_1 \ldots]] \ldots]\)
    b. \([\text{DP} \ldots [\text{NP}'] [\text{S'} [\text{D}]] \ldots]\)

Culy (1990:82-93) argues, contra Williamson (1987), that the relation between the outer and inner NP (or DP) is the same as the relation between antecedent and relative pronoun such as in English postnominal relative constructions. Three important common properties are (in my terms):

(50) a. There is \(\phi\)-feature agreement between \(\text{NP}_{\text{ant/out}}\) and \(\text{NP}_{\text{wh/in}}\).
    b. Restrictive relatives (adnominal or circumnominal) can be stacked, so \(\text{NP}_{\text{ant/out}}\) can relate to several \(\text{NP}_{\text{wh/in}}\).
    c. The relation between \(\text{NP}_{\text{ant/out}}\) and \(\text{NP}_{\text{wh/in}}\) is unbounded in principle, but there are constraints. In other words, the well-known characteristics of \(wh\)-movement show up in circumnominal relatives, too.

Especially because of (50a), the two NPs are co-indexed. (Notice, however, that (49a) is an \(i\)-within-\(i\) configuration.)

Like an English \(that\)-relative, a circumnominal relative contains a covert \(wh\)-element, which must be moved to the \(\text{COMP}\) domain at LF (cf. Culy 1990:95-99). Culy’s theory within a Government & Binding framework is given in (51).

(51) a. \(\text{D/S-structure}: [\text{NP} \ldots [\text{NP}'] [\text{S'} [\text{comp} [\text{S} \ldots [\text{NP} \ldots [\text{NP}_1 \ldots]] \ldots]] \ldots]] \ldots]\)
    b. \(\text{LF-structure}: [\text{NP} \ldots [\text{NP}'] [\text{comp} [\text{S} \ldots [\text{NP} \ldots [\text{NP}_1 \ldots]] \ldots]] \ldots]] \ldots]\)

This has the following advantages (in my terms):

(52) a. The \(wh\)-element mediates in the co-indexing relation between \(\text{NP}_\text{m}\) and \(\text{NP}_\text{out}\).
    b. The LF configuration with a \(wh\) in \(\text{COMP}\) generalizes over adnominal and circumnominal relative clause types.\(^{32}\)
    c. The LF configuration containing the \(wh\)-element is interpreted as a relative clause, contrary to other (nominalized) sentences.
    d. The general theory about \(wh\)-movement explains (50c): the constrained unboundedness.

### 5.2. Historical developments concerning the analysis

Before I continue with the incorporation of these ideas into the present framework, a brief summary of the historical development of the syntax of circumnominal relatives is in order. See Appendix III, Culy (1990:103-110) and Basilico (1996) for some more details.

Wilson (1963) derives circumnominals from adnominal relatives in an early transformational framework. This would involve lowering (of NP) in contemporary

\(^{32}\) Culy (1990:98) has a ‘Relative Co-indexing Constraint’ that generalizes over adnominal and circumnominal relatives.
terms. Gorbet (1976) and Hale & Platero (1974) represent circumnominal relatives as nominalized clauses, where nothing is moved. Platero (1974) and Weber (1983) represent them as sentences adjoined to an antecedent NP — i.e. there is an internal and an external NP — which is deleted. Peterson (1974) is perhaps the first who accommodates for an external determiner. Again, there is an internal and an external NP, where the latter is deleted.

Cole (1987) replaces deletion by covert LF movement of the internal to the external NP position. For similar ideas, see Broadwell (1985), Lefebvre & Muysken (1988) and Cole & Hermon (1994). Therefore at LF a circumnominal relative looks like an adnominal one. Unfortunately, Cole does not accommodate for the external determiner. Moreover, his account is based on crosslinguistic generalizations that turn out to be wrong. Therefore his approach is criticized in Culy (1990). Furthermore, Itō (1986), who discusses Japanese circumnominal relatives, has a theory comparable to Cole’s, except that — for Japanese in particular — the PF-derivation is distinct from the LF-derivation (see Appendix III).

Williamson’s (1987) theory on Lakota circumnominal relatives does have an external determiner. In addition, there is head raising of the internal NP to a position adjoined to Sʾrd at LF. Similar ideas can be found in Barss et al. (1990) and Bonneau (1992). The reason for this is — according to Williamson — that negative indefinites and irrealis determiners must be in the scope domain of a negative/irrealis marker, which is in the matrix clause. However, Culy (1990:182-197) has a more general semantic explanation for this phenomenon. I cannot repeat it here, since that would lead too far afield. Basilico (1996), too, argues that the head noun itself need not be raised out of its clause (see below for discussion).

Finally, Fontana (1989) argues that circumnominal relatives are like correlatives: they are supposed to be left-dislocated in the matrix sentence. The matrix sentence itself contains a pronoun, possibly zero. This is quite wrong: circumnominal relatives can be positioned at any argument position in the middle of the matrix clause. Moreover, there is no additional demonstrative, regularly. It is not even true for Lakota, because there every argument NP can get an additional demonstrative, optionally. The differences between correlative and circumnominal relatives are discussed further in the next section.

In short, I agree with Culy (1990) that i) circumnominal relatives are nominalized, ii) there is an external determiner position; iii) there is wh-movement; iv) circumnominal relatives are in several ways distinct from correlatives; v) there are generalizations covering the syntax and semantics of adnominal and circumnominal relatives.

5.3. The derivation of circumnominal relatives

At this point, consider how the present promotion theory of relative clauses applies to the syntax of circumnominal relatives in detail.

33 Culy (1990:254-259) criticizes hō’s treatment of no as a complementizer. According to Culy it is a nominalizing particle. Notice also that Murasugi (2000) claims that so-called circumnominal relatives in Japanese are misanalysed; they are not relatives at all. I cannot judge in this matter.
For once, the necessary movements to derive the word order appear to be simple. They are summarized in Table 3, based on Ch3§3.2.

**Table 3.** Movements in promotion theories required to derive the word order in circumnominal relative constructions.

<table>
<thead>
<tr>
<th>Subtheory</th>
<th>Antisymmetry, rigid left, func left-l.u.b.</th>
<th>Uniform branching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word order</td>
<td>VO and OV</td>
<td><strong>VO</strong></td>
</tr>
</tbody>
</table>
| Linear order | \( \text{spec } D \left[ \text{CP spec } \ldots \text{NP } \ldots \right] \) | \( \left[ \text{CP } \ldots \text{NP } \ldots \text{spec} \right] D \text{ spec} \)
| \( M_{\text{DP,}^{\phi}} \) | \( \checkmark \) |

Hence in an antisymmetric, rigid left or fl-lub promotion theory, circumnominal relatives are roughly like (53a) – to be extended below –, where the entire relative clause moves to SpecDP. This looks like a kind of pied piping again, such as discussed before. The uniform branching theory distinguishes between VO and OV languages. In the VO case the representation is like (53a), in the OV case, no movements are necessary, as in (53b), which is left-branching.

(53) a. \( [D_{\text{DP}} \left[ \text{CP } (C) \ldots \left[ D_{\text{rel}} \left( D_{\text{rel}} \right) [\text{NP } N] \ldots \right] \text{comp } (D) \right] t_{\text{CP}} \) \( \left( \text{spec-head-comp' base} \right) \)

b. \( [D_{\text{DP}} \left[ \text{CP } \ldots \left[ D_{\text{rel}} \left[ \text{NP } N \right] \left( D_{\text{rel}} \right) \ldots (C) \right] \right] (D) \) \( \left( \text{comp-head-spec' base} \right) \)

The head noun is in situ, hence there is no overt promotion. This means that the relative DP does not move to SpecCP. Thus suppose that the wh-feature is weak, which leads to (covert) feature movement.

First, consider the derivation of circumnominal relatives in OV languages in the uniform branching theory (53b). Since all constituents are lexically in the right position, all feature checking must be covert. Therefore after the internal checking in \( D_{\text{rel}} \) the formal features of \( D_{\text{rel}} \) move to C covertly for wh-checking. Next, they should be raised to the outer D for \( \phi \)-feature checking. However, that is not possible, since it would imply excorporation (of FF(\( D_{\text{rel}} \)) from C). Thus the uniform branching theory, although attractive at first sight in this respect, has a checking problem. I will show directly below that the other theories do not meet this problem.

Thus consider (53a) in detail. What is the status of the \( \phi \)-features of \( D/D_{\text{rel}} \)? Given property (48dii) I conclude that these are strong in some languages and weak in others. Namely, if in a VO language the normal order is D N, the \( \phi \)-features must be weak: there is no overt N(P)-movement. If the order is N D, there is overt movement, hence the features are strong.

The first possibility I want to explore is: wh weak and \( \phi \) strong on a spec-head-comp base. Initially, N incorporates into \( D_{\text{rel}} \) (or NP moves to SpecDPrel), so that N and \( D_{\text{rel}} \)'s Case and \( \phi \)-features can be checked overtly. \( D_{\text{rel}} \) does not need to move to SpecCP overtly, since wh is weak. Instead, \( D_{\text{rel}} \)'s formal features move to C and wh is checked covertly. CP is selected by the matrix clause D. D's \( \phi \)-features are strong hence something must be attracted. \( D_{\text{rel}} \)'s features are present in C, hence in CP, so what happens is that CP is moved to SpecDP, where the \( \phi \)-features can be
checked in spec-head configuration. This derivation is indicated in (54) in some detail.

\[(\text{DP} [\text{CP} [C \text{FF(D$_{rel}$)+C)] [\text{IP} \ldots [\text{DP$_{rel}$} [D$_{rel}$ N+(D$_{rel}$)] [\text{NP}_{CP}] \ldots ]]\text{CP} (\text{D}) t_{CP}]\]

Notice that this derivation explains (48e): the external D is final. It may also explain
(48ci): there is no relative pronoun, i.e. D$_{rel}$ must be covert. At least a part of D$_{rel}$
c-commands N, hence there is a threat of a Binding Principle C violation.

Some unwanted possibilities must be excluded. First, the derivation crashes if
only D$_{rel}$'s wh-feature moves to C, because there are no available $\phi$-features for
the matrix D to check with. So all formal features of D$_{rel}$ are pied piped to C,
although only wh is attracted. This is in accordance with standard assumptions.
Second, why doesn't FF(D$_{rel}$) move on to D, instead of moving the whole CP to
SpecDP? This is because i) exorporation (here, of D$_{rel}$ from C) is not possible in
general; and ii) if the Cases of D$_{rel}$ and D are different, incorporation of D$_{rel}$ into D
leads to a crash. Third, instead of CP movement, why does C – or more precisely
[c FF(D)+C] – not move to D? I am not sure how to exclude this technically, but I
cannot even think of a possible interpretation of incorporation of a complementizer
into a determiner of a higher clause. (Whereas the other way round, movement of an
argument to the CP domain is a way of scope-marking.) So suppose this is not an
option. Then (54) is the only possible derivation given this feature setting, which
is what is desired.

Next, consider what happens if the $\phi$-features are weak, i.e. in the D N
languages. This derivation only minimally differs from (54). It is shown in (55).

\[(\text{DP} [\text{CP} [C \text{FF(D$_{rel}$)+C)] [\text{IP} \ldots [\text{DP$_{rel}$} \text{FF(N)+}(D_{rel})] [\text{NP}_{CP}] \ldots ]]\text{CP} (\text{D}) t_{CP}]\]

The head noun moves covertly to D$_{rel}$, since the features are weak. As in (54), D$_{rel}$
moves to C covertly. I have argued that exorporation of D$_{rel}$ is impossible and that
C cannot incorporate into the matrix D. (Whether this would be overt or covert is
irrelevant.) Thus, even if $\phi$ is weak, CP must move to SpecDP in order to prevent a

So even in D N languages, a determiner must be final in relative constructions.
Perhaps this explains why it is never pronounced in these languages, at least not in
those described in the data set, since an overt D in these relative constructions
would go against the normal pattern. In other words: the syntax of circumnominal
relative constructions forces a determiner to be construction-final in all relevant
languages. In those with a D N pattern, this is odd from a perceptual point of view,
which may be the reason why an overt D is shunned in relative constructions.

---

34 See Culy (1990:261) on Diegueño, Dogon, Lakota, Japanese, Navaho, the Quechua family, Moore,
Dagbani and ASL. Four additional languages, Crow, Mohave, Tibetan and Yavapai (from Lehnmann
1984) conform to this pattern.

35 This may also be the explanation for the fact that circumnominal relatives generally do not occur in
D N languages where D is regularly overt, since in that case every possible relative would have to be
shunned, which makes the relative strategy vacuous.
5.4. Additional issues

Some aspects of circumnominal relatives deserve further discussion. I would like to add some notes on the indefiniteness effect, verb morphology, nominalization and cross-linguistic generalizations, internal head movement, island effects, and maximalization.

5.4.1. The indefiniteness effect

Williamson (1987) describes an indefiniteness effect for the internal head in Lakhota circumnominal relatives. It is confirmed by Culy (1990) for other languages. According to Culy there are several semantic explanations for this phenomenon. The most convincing one, I believe, is the following. DP's are generalized quantifiers, hence the circumnominal relative must contain a free variable to be bound by D. If the head noun phrase is a definite DP, the quantification by the outer D would be vacuous, so this is excluded. I will not further discuss the indefiniteness effect, but see Williamson (1987) and especially Culy (1990:Ch3), who also notes and explains some counterexamples.

5.4.2. Verbs and morphology

Culy (1990:128-150) discusses some morphological issues concerning nominalized sentences. Since these do not concern circumnominal relatives in particular, but also adnominal relatives and other nominalized sentences, I will not discuss them here, but simply list his conclusions:

... continued

[36] Provided that the proposed syntax is correct, this functional claim is much stronger and more straightforward than Culy's (1990:207-239) proposal which makes use of the Consecutive/Embedding Constraint (CEC) which states that centre embedding in combination with a consecutive order of equal lexical elements is ruled out. For instance, the CEC predicts that \([np \text{ Det} [n [\text{ Det N}] \text{ (Obj, V) }]]\) is impossible, but the simple addition of an adverb preceding the subject should annul the effect, which is not true. Similarly, if the two adjacent determiners are different, there is no violation. Therefore Culy proposes that configurations that potentially violate the CEC can be generalized to be ungrammatical. However, this generalization incorrectly includes SVO, N Det languages, which produce a potential CEC violating configuration. So Culy assumes that there is language variation on this point. My proposal does not have these problems, although much more detailed data are necessary for a good comprehension of the issue. (To mention just one difficulty: Culy claims on the basis of positive evidence only, that indefinite IHRCs like \([n \text{ Det} [n [\text{ Det N} \text{ indef}] \text{ indef}]]\) do not occur in the SVO, N Det language Moore, whereas definite ones do: \([n \text{ Det} [n [\text{ Det N} \text{ indef}] \text{ def}]]\). This follows from the CEC. However, Culy seems to overlook that again an adverb - here sentence-final - would bypass the CEC effect. Hence more data, including negative judgements are necessary.) Still, I do not wish to claim that the CEC (or an equivalent functional filter) in the strict sense is not a real effect. It just seems not strong enough to explain the distribution of determiners in circumnominal relative constructions.

[37] See also the section on relative affixes in Ch5§4.1. For a list of relevant morphological characteristics of several circumnominal strategies, see Culy (1990:262).
5.4.3. **Nominalization and cross-linguistic generalizations**

In earlier work, e.g. Cole (1987) and Downing (1978), it is assumed that circumnominal relatives occur in languages with characteristics such as: SOV order and *pro* drop. Further research – in particular Culy (1990:Ch4) – has shown that neither is true; cf. (48) above, and – concerning *pro* drop – Culy (1990:240-242). Instead, it turns out that the presence of nominalized sentences is crucial. Culy (1990:203) states: “A language will have circumnominal relatives only if it also has other [read: at least one type of, MdV] similar nominalized sentences with the independency properties [i.e. independent reference of arguments, and independent tense, mood and aspect].” Examples of other clauses that can be nominalized are factive complements, indirect questions, complements of verbs of saying, etc. Culy (1990:264) lists for several circumnominal relative languages which other nominalized sentence types they have.\(^{38}\)

5.4.4. **Internal head movement**

Circumnominal relatives have an internal head which is not *wh*-moved. However, according to Lehmann (1984:121), there are circumnominal relatives with a fronted head. As a main strategy this is only found in Yavapai (cf. Appendix II, table 3), but it is a secondary strategy in Gaididj, Mohave, Diegueño, Latin and Sanskrit. This phenomenon has remained unnoticed by Culy (1990). Lehmann argues that it may be compared with *attractio inversa*, where the antecedent of a postnominal relative gets subordinate clause Case. Perhaps this process marks the transition between the postnominal and the circumnominal strategy. If so, one might expect a link between circumnominal and prenominal relatives, too, namely by extraposition of the head within a circumnominal relative. Lehmann (1984:122, cf. 129) suggests that this may exist indeed in some (stylistically marked) literary variants in ancient Greek and Latin. There is a problem, however: there is a sentence-initial relative pronoun in the examples provided by Lehmann. Hence it looks like stranding of the antecedent in a left-oriented structure, which is therefore a ‘failed’ postnominal relative hence a circumnominal one. Thus this is another instance of the transition between circumnominal and postnominal relatives. (Moreover, it is another indication for a promotion theory of relatives.) A transition between prenominal and circumnominal relatives is not found, if it exists at all. Notice that *attractio inversa* is also not found in prenominal relative constructions, as far as I know.

Basilico (1996) discusses internal head movement in circumnominal relatives in more detail. He shows that there are two types of internal head movement:

\(^{38}\) There is one exception: ASL has no other nominalized clause types. Perhaps this is explained by the fact that ASL is the only IHRC language with a relative marker; cf. Culy (1990:205-206).
movement to the front of the sentence and movement to an intermediate position. There are examples from the Yuman languages (Mohave, Cocopa, and Diegueño), Northern Athapaskan (Tanaina, Koyukon), and Gur (e.g. Moore). The effect of such a movement is that the relative clause is disambiguated (in case that there are more arguments that could be the head). In all examples, internal head movement is optional. Basilico argues that it is necessary (whether overt or covert) in order to avoid existential closure (the relevant NPs are specific). Hence it is comparable with scrambling and object shift in Germanic languages. Moreover, the internal variable of the indefinite noun needs to be bound locally by the external determiner. This is in line with Williamson (1987), Culy (1990) and Srivastav (1991). I will not further discuss this issue, but see Appendix III for some details of Basilico’s approach.

5.4.5. Island effects

Circumnominal relatives, like adnominal relatives, are in principle unbounded, i.e. constructions of the type the man whom I thought that you saw are attested. If wh-movement is involved, one would expect that island effects play a role. There is at least some information concerning violations of the Complex Noun Phrase Constraint (both question word extraction and ‘boundedness at LF’, i.e. (covert) wh-movement in relative clauses), the Coordinate Structure Constraint, and the Empty Category Principle (subject/object asymmetry). The relevant construction types are illustrated with well-known English examples in (57).

(57) a. CNPC (question extraction): * who do you love the child that saw yesterday?  
   b. CNPC (boundedness at LF): * the newspaper that we talk to many people who read is the Times.  
   c. CSC: * the dog which and the cat were fighting is barking.  
   d. ECP: * the man who I thought that saw you.  

the man who I thought that you saw _.


---

39 The ECP effect mentioned is known as the that-trace effect. In this light, the effect is a bit strange here, since circumnominal relatives do not contain complementizers. The English examples without the complementizer that do not display the effect: both are good. In Dutch, the complementizer cannot be omitted, but still both examples are equally questionable for me – but equally good for some others, see e.g. Bennis (1986).
As expected, there is some language variation, which is well-known from adnominal relative constructions. Culy stresses that, as far as can be seen, the effects are equal for adnominal and circumnominal relatives within one language. (Recall that most circumnominal relative languages also have adnominal RCs.) Of course they should also equal the effects in non-relative contexts. Notice that Japanese and Lakota seem to be quite liberal. If they do not obey any constraint whatsoever, this could be problematic for the theory of wh-movement. However, I do not expect so. Clearly much more data are needed on this issue. Hopefully, the results of a systematic investigation into island effects in a larger number of languages will be available some day.

5.4.6. Maximalization and subjacency

Grosu & Landman (1998) suggest that circumnominal relatives without an overt determiner are maximalizing (cf. Ch2§3 on Grosu & Landman’s scale). If determiners are overt, the interpretation is restrictive. This hypothesis is based on Quechua and Lakota. It is not clear if this is cross-linguistically so, and why. Basilico (1996:518) predicts that if there are subjacency violations, stacking is possible. In order to check these two hypotheses, I have collected the following table. The data are from Cole (1987), Itô (1986), Williamson (1987), Culy (1990:110-128, 214-223, 261), Basilico (1996), and Grosu & Landman (1998). Recall that maximalization is indicated by (external) determiner restrictions (only definite and universal) and a prohibition of stacking.

### Table 4. Island effects in circumnominal relatives.

<table>
<thead>
<tr>
<th>language</th>
<th>unboundedness</th>
<th>CNPC(LF) obeyed</th>
<th>CNPC(qu) obeyed</th>
<th>CSC obeyed</th>
<th>ECP effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quechua</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Japanese</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>?</td>
<td>yes</td>
</tr>
<tr>
<td>Lakota</td>
<td>yes</td>
<td>no</td>
<td>?</td>
<td>?</td>
<td>no</td>
</tr>
<tr>
<td>Mohave</td>
<td>?</td>
<td>no</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

### Table 5. Some characteristics in particular circumnominal relative strategies

<table>
<thead>
<tr>
<th>language</th>
<th>island violations (at least CNPC(LF) )</th>
<th>overt D</th>
<th>stacking</th>
<th>determiner restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quechua</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Navaho</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>?</td>
</tr>
<tr>
<td>Japanese</td>
<td>yes</td>
<td>no</td>
<td>no?&lt;sup&gt;40&lt;/sup&gt;</td>
<td>?</td>
</tr>
<tr>
<td>Lakota</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Mohave</td>
<td>yes</td>
<td>yes</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Moore</td>
<td>?</td>
<td>yes</td>
<td>?</td>
<td>yes</td>
</tr>
<tr>
<td>Dagbani</td>
<td>?</td>
<td>yes</td>
<td>?</td>
<td>no</td>
</tr>
</tbody>
</table>

<sup>40</sup> Strangely, Japanese is reported to have multiple headed relatives (cf. Itô 1986).
If the data on Moore (from Culy 1990) are correct, they disprove Grosu and Landman's hypothesis partly. However, it may still be correct that the absence of a determiner forces a maximalizing reading. Second, if Japanese disallows stacking (a statement from Grosu & Landman 1998), Basilico's claim is incorrect. Nevertheless, more data are necessary before definitive conclusions can be drawn.

6. Correlations

The fourth syntactic main type of relative constructions is the correlative one. It is discussed in three subsections: 6.1 is an overview of the properties of correlatives; 6.2 discusses briefly the history of the analysis; and 6.3 presents the derivation within the present framework.

6.1. Introduction and properties

Correlatives occur in various language families across the world (cf. Appendix II, figure 1 and table 4). Usually, they have a structure like (58), where the relative CP is left-adjointed to the matrix clause, the head noun is internal to the relative clause, and the matrix contains a demonstrative correlate.

(58) \[ \text{matrix} \left[ \text{CP} \right. \left. \left[ \text{DP-rel \ wh} \ NP \right] \ldots \right] \ldots \text{matrix} \ldots \text{Dem} \ldots \] \]

Correlatives are preposed co-relatives (cf. the terminological chart in Ch2§2.5). In various publications (right-)extraposed relatives are also called correlatives, but I will not use this confusing terminology. Srivastav (1991) shows that extraposed relatives are clearly distinct from preposed correlatives; see below. Extraposed relatives behave like adnominal relatives. Thus I will refer exclusively to structures like (58) as correlatives. (Extraposition is treated separately in Ch7.)

Correlatives differ from circumnominal relatives (although on the basis of an individual sentence in a particular language the distinction may be hard to make); see table 4.
Table 4. Differences between correlatives and circumnominal relative clauses.

<table>
<thead>
<tr>
<th>property</th>
<th>correlative</th>
<th>circumnominal</th>
</tr>
</thead>
<tbody>
<tr>
<td>relative pronoun in RC</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>pronoun in matrix</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>RC is sentence-initial (in matrix)</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>RC is in a DP position (in matrix)</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>category of RC construction</td>
<td>CP</td>
<td>DP</td>
</tr>
<tr>
<td>RC construction is nominalized</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>external determiner</td>
<td>no</td>
<td>possibly</td>
</tr>
<tr>
<td>external Case marking</td>
<td>no</td>
<td>possibly</td>
</tr>
<tr>
<td>external adposition</td>
<td>no</td>
<td>possibly</td>
</tr>
<tr>
<td>indefiniteness restriction on head noun</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>RC is maximalizing</td>
<td>yes</td>
<td>not necessarily</td>
</tr>
</tbody>
</table>

Hence I agree with Culy (1990:26) that correlatives and circumnominals are separate phenomena, although of course both are relative constructions, and there are obvious and not-so-obvious similarities (cf. Lehmann 1984; Srivastav 1991 and Grosu & Landman 1998).

5.2. Historical developments concerning the analysis

Consider briefly the history of the analysis of correlatives (see Appendix III for some more details). One possible view is that they are generated as adnominal relatives and subsequently moved to a position left-adjointed to the matrix; see e.g. Verma (1966), Junghare (1973), Kachru (1973, 1978), Wali (1982), or Subbarao (1984). Another view is that correlatives are syntactically different from English-type relatives: they are generated in a left-peripheral position; cf. (58). This is advocated in Donaldson (1971), Downing (1973), Bach & Cooper (1978), Dasgupta (1980), Lehmann (1984), Keenan (1985), and Andrews (1985).

The first approach assumes that both the syntax and the semantics of adnominal and correlative relatives is basically the same. The second approach has a different syntax, but a similar semantics (cf. Srivastav 1991 for details). More recent research has shown that correlatives differ from adnominal relatives in important respects. Concerning the semantics, Srivastav (1991) argues that correlatives are quantificational expressions. Grosu & Landman (1998) extend her ideas into a more general perspective, in which correlatives are one of several construction types that are semantically maximalizing. Hence the demonstrative correlate in the matrix clause is definite or universal only. Moreover, stacking is impossible.45,46 See e.g. the Hindi examples in (59) and (60), taken from Grosu & Landman (1998:164/5).

41 But possibly null.
42 Except if the language allows for doubling pronouns, as in Lakota.
43 Unless accidentally, or if it is scrambled.
44 Cf. Ch2§3 on Grosu & Landman’s scale.
A further difference is the presence or absence of the head noun in the internal and external position. This can be shown nicely in Hindi, which has correlative, adnominal and extraposed relatives. It turns out that extraposed relatives behave on a par with adnominal relatives. Concerning the head noun, Srivastava (1991) shows data that can be schematically summarized as follows:

(61) a. [\textit{wh} \textit{N} \ldots] \ldots \textit{Dem N} \ldots [correlative]
   b. [\textit{wh} \textit{N} \ldots] \ldots \textit{Dem} \ldots
   c. [\textit{wh} \ldots] \ldots \textit{Dem N} \ldots

(62) a. * \ldots \textit{Dem N} [\textit{wh} \textit{N} \ldots] \ldots [postnominal/extraposed]
   b. * \ldots \textit{Dem} [\textit{wh} \textit{N} \ldots] \ldots
   c. \ldots \textit{Dem N} [\textit{wh} \ldots] \ldots

In other words, postnominal relatives display an antecedent head noun only. In correlatives, there is possible variation: the head noun is either internal or external or both.

A third difference is that correlatives allow for multiple relativization such as 
\textit{jis laRkiiNe jis laRkeKO dekhaa usNE usKO pasand kiyaa} ‘Which girl saw which boy, she liked him.’ This kind of sentences are impossible to construct with adnominal or extraposed relatives. For those reasons, Srivastava (1991) proposes the following structure for correlatives:

(63) \textit{IP} [\textit{CP} \textit{wh N} \ldots], [\textit{IP} \ldots \textit{Demi} \ldots]]

\textit{continued}

46 If I understand Bianchi (1999:90) correctly, recursive embedding as in \textit{I saw the boy who saw the girl who saw you} is also impossible in correlative constructions. Clearly, it would lead to centre embedding, but I fail to see how it is excluded exactly. Bianchi (who refers to Srivastava) states that material intervenes between the correlative sentence and the variable in the matrix clause. This is indeed the case linearly, but not hierarchically.

47 In fact, there are three options (cf. Ch2§7.6): (i) with bijection: [\textit{wh1 wh2 \ldots Demi Demj}], (ii) with a split correlate: [\textit{wh\textsubscript{parent} \ldots Demi Demj}], and (iii) with a split wh: [\textit{wh1 wh2 \ldots Dem\textsubscript{plural}}], see Grosu & Landman (1998) for examples. Notice that the separated \textit{wh} and/or \textit{Dem} phrases can have different roles in the relative clause and matrix clause, respectively. Thus these are quite different from the rare examples with a split antecedent that can be found in languages with another relative strategy, e.g. \textit{English a man, went out and a woman, came in who\textsubscript{i}, were engaged}. In this kind of examples coordination and role equivalency is crucial. Impossible is for instance: \textit{a man, saw a woman, who\textsubscript{i}, were walking}. 
Here the relative CP is left-adjointed to the matrix IP. This CP is a quantifier that binds the demonstrative in the matrix clause. The relative is adjoined to IP, since if the whole construction is embedded (e.g. in a factive context), it follows the complementizer. Hence the correlative cannot be in SpecCP or adjoined to CP.

If the head noun’s number differs from Dem, the verb in the relative clause agrees with N, hence the relative is not in the scope of the matrix demonstrative. Srivastav’s example is Jo laRke khaRe hāī har ek meraa chaatr hai ‘Which boys are standing, each one is my student.’ Again, this indicates that the correlative is base-generated in a left-adjointed position.

As for the internal syntax in Hindi correlatives, the relative wh-operator need not be moved overtly to SpecCP<sub>rel</sub>. This corresponds to the fact that wh-questions are in situ. Strange, however, is the optionality involved.

Finally, the demonstrative in the matrix clause is like a variable. The distance between the correlative and the demonstrative cannot be too large: “Dem is a locally A’-bound pronominal” (Srivastav 1991:680). Grosu & Landman (1998:167) propose an improvement on this by stating that “we do not assume the correlate [= the demonstrative] itself to be a variable, interpreted in situ: there is a variable bound by abstraction in the position of the correlate, but the meaning of the correlate itself contributes to the building of a generalized quantifier outside the IP”.

5.3. The derivation of correlative constructions

At this point consider how the correlative construction relates to the promotion theory of relative clauses. Essential facts for the syntax are the following ones:

(64) a. There is no ‘external determiner’: the correlative is a CP.
   b. Usually there is a relative pronoun bearing subCase in the correlative.
   c. The internal head noun accompanying D<sub>rel</sub> (if present) bears subordinate clause Case.
   d. Usually there is overt wh-movement in the correlative.
   e. The matrix demonstrative (the correlate) bears matrix Case.
   f. If there is a resumptive head noun – as in (61a/c) –, it bears matrix Case.

The grammar of the matrix clause is rather obvious. There is an argument position that is syntactically filled by a DP which is i) zero, if the language allows for pro drop, or ii) a (definite or universal) demonstrative/personal pronoun, or iii) a demonstrative plus a resumptive head noun (where of course this demonstrative must be usable as a dependent D). This DP checks Case in the matrix clause.

The correlative clause is adjoined at some point, probably IP (I will return to this). Since there is no external determiner, there is no trigger for head raising. Thus consider the feature checking internal to CP<sub>rel</sub>, e.g. which girls you saw, to be concrete. The relative DP which girls is generated in the direct object position. D and N agree in φ-features and bear the same Case. If this were not the case, the derivation would crash, since there is no external determiner to rescue it. N’s features need to be checked, hence the formal features of N incorporate into D<sub>rel</sub> – the most economical solution. DP<sub>rel</sub>’s accusative Case is checked with AgrO.
Finally, the *wh*-feature present (usually strong) triggers movement to SpecCP. This is indicated in (65).

\[
(65) \left[ \text{CP-corr} \left[ \text{DP-rel} \left[ D \text{FF(N)+D}_{\text{rel}} \right] \text{[NP N]} \right] \left[ \text{C} \left[ \text{IP DP}_{\text{subj}} \left[ \text{AgrOP t} \right] \text{AgrO} \left[ \text{VP t}_{\text{subj}} \text{V t} \right] \right] \right] \right] \text{ which girls you saw}
\]

I conclude that the syntactic derivation of correlative is straightforward.

Bianchi (1999:86-88) notes that many languages with more than one relative strategy (among which Hindi) use the same relative pronouns in postnominal and correlative clauses. From the perspective of the standard theory of relatives this is not necessarily the case: in a correlative clause DP_{rel} is *-[which girls]* (i.e., D_{rel} is a determiner selecting the head noun); in a postnominal relative DP_{rel} is just *-[which]* (i.e., D_{rel} is an independent pronoun anaphoric to an antecedent). By contrast, in the promotion theory DP_{rel} is equal in both cases.

Finally, the position of CP_{corr} in the matrix clause must be considered somewhat more precisely. Potential base positions for CP are:

\[
\begin{align*}
\text{(66) a.} & \quad \text{the complement of D}_{\text{dem}} \\
\text{b.} & \quad \text{SpecDP}_{\text{dem}} \\
\text{c.} & \quad \text{AdjDP}_{\text{dem}} \\
\text{d.} & \quad \text{AdjIP} \\
\text{e.} & \quad \text{SpecCP} \\
\text{f.} & \quad \text{AdjCP}
\end{align*}
\]

Option (66a) can be excluded immediately—even apart from Srivastav’s semantic arguments. CP cannot take the base position of an adnominal relative since that may already be filled with a resumptive head noun, as indicated before; cf. (61a/c).

SpecCP and AdjCP (66e/f) are excluded, because a correlative follows a complementizer if the whole construction is embedded, as noted before. SpecDP (66b) and AdjDP (66c) are excluded because of the multiple relativization cases. (If there are two correlates, to which DP_{dem} should CP_{corr} be connected?) Thus I agree with Srivastav and others that AdjIP (66d) is a plausible base position.

---

48 Generally, AdjIP is a position open for extra material. For instance, temporal adverbs and adverbial clauses can also be generated there. This can be shown in Dutch, where, in main clauses, the subject and the finite verb are in SpecCP and C respectively, and a definite object is scrambled out of VP, say, to SpecAgrOP: *Ik heb [gisteren / toen oma kwam] de hond met plezier uitgelaten* [I have *yesterday / when grandma came* the dog with pleasure taken out].

49 However, there are examples where the correlative seems to be adjoined to the demonstrative DP, i.e. in the middle of the matrix clause, cf. Wali (1982) and Srivastav (1991). So AdjDP (66c) is a position for correlative in exceptional cases. (It is not clear to me whether SpecDP (66b) is a possible alternative for this.) But then one could also analyse a correlative in AdjIP as if it has moved there from AdjDP (or SpecDP). Neither Srivastav nor Bianchi considers this option, probably because of the multiple relativization cases. However, these in turn could be exceptional. It seems to me that it has certain advantages to take AdjDP (or SpecDP) as the basis. For instance, it would explain the locality effects between CP_{corr} and D_{dem} mentioned in a direct way. Preposing of the correlative could then be seen as a movement driven by information structural preferences, as is the case for other types of scrambling. Furthermore, it allows us to assume one base position for all...
One final remark is in order. In the strict version of antisymmetry, adjunction is not possible, which is problematic for correlatives. (Strikingly, this is ignored by Bianchi 1999.) Hence again there are three options: i) a more elaborate phrase structure is needed in order to create a position for correlative clauses; ii) Zwart’s (1993) revision of antisymmetry is used, which accounts for one adjoined position; or iii) strict antisymmetry is given up for a rigid left, uniform branching or fl-lub version of phrase structure.

7. The syntax of main types of relatives: summary and conclusion

This chapter has treated of the promotion theory of relativization in detail. I have briefly commented upon Kayne’s (1994) and Bianchi’s (1999) version of it and upon Borsley’s (1997) critique. The derivations of all syntactic main types of relatives have been discussed. I have argued that all types of relatives involve the same ‘ingredients’, and that the differences can be traced back to overt/covert distinctions, that is, differences in the feature checking procedure. The only features that are relevant in this respect are *wh*, Case, and $\phi$-features. Furthermore, there are differences in pied piping, which is accounted for in terms of covert feature movement. Namely, if formal features move up without directly establishing a checking relation, i.e. ‘percolation’, this causes pied piping. Finally, I have concluded that the uniform branching theory of phrase structure is untenable from the perspective of derivations of relative constructions. Rather, a universal spec-head-comp basis is preferable, at least in the functional domain, which is what is relevant in a relative context.

I will provide an overview of all analyses here, and discuss which ‘parameter setting’ leads to which type of relative construction. Obviously I cannot repeat all potential alternatives that lead to a crash, but see the text above.

Consider the features involved. First, there is the *wh*-feature (on C and D$_{rel}$). If it is strong then D$_{rel}$ moves to SpecCP; if it is weak, only the formal features (FF) of D$_{rel}$ move to C, which gives a circumnominal relative. This FF movement causes heavy pied piping (HPP) of the relative CP (CP$_{rel}$) to the specifier of the external (matrix) determiner (SpecDP$_{ext}$); see below. Second, Case is important. The relative DP (DP$_{rel}$) checks subordinate clause Case (subCase) with some X in the

... continued

correlatives (except for the multiple relativization cases). If so, this reflects in a direct way that CP$_{corel}$ and D$_{dem}$ (+N) are semantically interpreted as a unit. Still, as required, D$_{dem}$ does not c-command the relative clause (unlike in adnominal relative constructions).

So far, I have ignored the following problem: there are languages which seem to have a contradictory *wh*-feature in relative clauses and questions. For instance, in Imbabura there is overt *wh*-movement in questions, but relatives are circumnominal, in which *wh* is weak. Chinese shows the complementary pattern: it has *wh*-in-situ questions, but prenominal relatives, in which *wh* is supposed to be strong. At present I am not sure how to proceed on this matter.
relative clause. In turn, the ‘external’ DP (DP$_{ex}$) in the matrix clause (CP$_{mat}$) checks matrix clause Case (matrCase) with some X in the matrix, whether overt or covert. Furthermore, the relative N must check matrCase with D$_{ex}$. Third, the $\phi$-features (on nouns and determiners) have their influence. Those on N are checked with DP$_{rel}$. If $\phi$ on D$_{ex/rel}$ is strong, then N moves to D (or NP to SpecDP). If it is weak, then FF(N) moves to D, or, in the relative clause, NP moves to SpecDP$_{rel}$ (except in correlatives and circumnominal relatives).

In this context heavy percolation is feature movement of N to C$_{rel}$. It is only useful if D$_{ex}$ has strong $\phi$-features, which then causes heavy pied piping for $\phi$-feature checking. I have argued that HPP must be CP$_{rel}$ movement to SpecDP$_{ext}$, and not IP movement. There are two types of percolation here. Type A is full FF(N) percolation to C; type B is percolation of the $\phi$-features only. In the first case N’s Case feature is checked as well via CP; in the second this is not possible, hence in addition to HPP, N must move to D.

Table 5 lists the different settings for all relative clause types. Notice that the position of the correlative with respect to the matrix is different from the other types of relatives, but the features in use are the same everywhere.

---

**Table 5. Parameter settings for different relative clause types in a rigid-left promotion theory.**

<table>
<thead>
<tr>
<th>RC type</th>
<th>wh</th>
<th>$\phi$ (on D)</th>
<th>heavy percolation</th>
<th>remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>post</td>
<td>D N RC</td>
<td>strong</td>
<td>weak</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>N D RC</td>
<td>strong</td>
<td>strong</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>N RC D</td>
<td>strong</td>
<td>strong</td>
<td>type A</td>
</tr>
<tr>
<td>pre</td>
<td>D RC N</td>
<td>strong</td>
<td>weak</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>RC D N</td>
<td>strong</td>
<td>strong</td>
<td>type B</td>
</tr>
<tr>
<td></td>
<td>RC N D</td>
<td>strong</td>
<td>strong</td>
<td>type B</td>
</tr>
<tr>
<td>cir</td>
<td>RC D</td>
<td>weak (irrelevant)</td>
<td>no</td>
<td>HPP</td>
</tr>
<tr>
<td>cor</td>
<td>RC…Dem</td>
<td>(strong)</td>
<td>weak</td>
<td>no</td>
</tr>
</tbody>
</table>

These settings lead to the representations in table 7, which are derived on the basis of the steps indicated in table 6.

---

Notice that if wh is strong, this movement for Case checking must be overt, whether X is strong or weak. This is similar in normal questions. Rather obviously, the strong feature of the higher head blocks the more economical option which would be available for Case in another context.
Table 6. Movements in different relative clause types in a rigid-left promotion theory.

<table>
<thead>
<tr>
<th>RC type</th>
<th>movements</th>
<th>reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>post D N RC</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (weak) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF(N) → D_{ext}</td>
<td>D_{ext} checks ϕ (weak) &amp; matrCase with N</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>post N D RC</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (strong) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>N → D_{ext}</td>
<td>D_{ext} checks ϕ (strong) &amp; matrCase with N</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>post N RC D</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (strong) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF[N] → C(P)_{mat}</td>
<td>ϕ &amp; Case percolation → heavy pied piping</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>CP_{rel} → SpecDP_{ext}</td>
<td>D_{ext} checks ϕ (strong) &amp; matrCase with CP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>pre D RC N</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (weak) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>IP_{rel} → AdjCP_{rel}</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>FF[N] → D_{ext}</td>
<td>D_{ext} checks ϕ (weak) &amp; matrCase with N</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>pre RC D N</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (strong) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF[N] → C(P)_{mat}</td>
<td>ϕ percolation (not Case) → heavy pied piping</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>N → D_{ext}</td>
<td>D_{ext} checks matrCase with N</td>
</tr>
<tr>
<td></td>
<td>CP_{rel} → SpecDP_{ext}</td>
<td>D_{ext} checks ϕ (strong) with CP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>pre RC N D</td>
<td>NP → SpecDP_{rel}</td>
<td>D_{rel} checks ϕ (strong) with NP</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF[N] → C(P)_{mat}</td>
<td>ϕ percolation (not Case) → heavy pied piping</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>N → D_{ext}</td>
<td>D_{ext} checks matrCase with N</td>
</tr>
<tr>
<td></td>
<td>CP_{rel} → SpecDP_{ext}</td>
<td>D_{ext} checks ϕ (strong) with CP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>cir RC D</td>
<td>N → D_{rel} or FF[N] → D_{rel}</td>
<td>D_{rel} checks ϕ (strong or weak) &amp; subCase with N</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF[D_{rel}] → C_{rel}</td>
<td>C_{rel} checks ϕ with D_{rel}</td>
</tr>
<tr>
<td></td>
<td>CP_{rel} → SpecDP_{ext}</td>
<td>D_{ext} checks ϕ with CP_{rel} (heavy pied piping)</td>
</tr>
<tr>
<td></td>
<td>DP_{ext} → Case position in CP_{mat}</td>
<td>some X checks matrCase with DP_{ext}</td>
</tr>
<tr>
<td>cor RC...Dem</td>
<td>in the matrix:</td>
<td>some X checks matrCase with DP_{dem}</td>
</tr>
<tr>
<td></td>
<td>DP_{dem} → Case position in CP_{mat}</td>
<td>D_{rel} checks ϕ &amp; subCase with N</td>
</tr>
<tr>
<td></td>
<td>in the correlative clause:</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>FF[N] → D_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → Case position in CP_{rel}</td>
<td>some X checks subCase with DP_{rel}</td>
</tr>
<tr>
<td></td>
<td>DP_{rel} → SpecCP_{rel}</td>
<td>C_{rel} checks ϕ (strong) with DP_{rel}</td>
</tr>
</tbody>
</table>
Table 7. Structural representations for different relative clause types in a rigid-left promotion theory.

<table>
<thead>
<tr>
<th>RC type</th>
<th>Structural representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>post</td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td>pre</td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td>cor</td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Structural representation formula" /></td>
</tr>
</tbody>
</table>

In short, I have tried to provide a coherent and complete system to describe the syntactic aspects of relativization. In the body of this and the previous chapter I have excluded many potential alternatives. Of course I am aware that some details of the present analysis may not be completely satisfactory (especially concerning prenominal relatives), and they will probably be reconsidered in future research. Nevertheless, the promotion theory as argued for is the least implausible analysis of relativization in general.