Generalized Transformations and the Wh- cycle: free relatives as bare Wh-CPs

Johan Rooryck
Leiden University

In this paper, it is claimed that free relatives (FRs) are bare Wh-CPs, and as such syntactically equivalent to indirect Wh- clauses. Exploiting the power of Chomsky's (1992) Generalized Transformations (GT), this analysis thus challenges the notion that FRs involve a relative clause, either as a predicate of a null head (Kuroda 1968, Groos & Van Riemsdijk 1978, Hirschbühler 1978), or as an adjunct to a Wh-XP (Bresnan & Grimshaw 1978, Larson 1987). It will be shown that these traditional analyses of FRs do not satisfactorily explain the syntactic and semantic similarities between FRs and indirect Wh- clauses. At first sight, important differences between FRs and indirect Wh- clauses preclude an analysis of FRs and indirect Wh- clauses as bare Wh-CPs. The differences involve extraction, Matching, and the restriction of FRs but not indirect Wh- clauses to Wh- NPs and APs (Larson 1987). It will be argued that GT can adequately account for these differences if it is accepted that GT can insert both phrase-markers that have internally been subject to movement before GT insertion, and phrase-markers in which movement has yet to take place after GT insertion. The former situation gives rise to FRs, the latter to indirect Wh- clauses. Further differences between indirect Wh- clauses and FRs will be shown to derive from an economy principle which prevents a second application of the syntax to GT-inserted CPs which have internally undergone movement. It will be argued that this principle is independently motivated. If this analysis is on the right track, FRs provide important insights into the nature of economy principles in the syntax, in keeping with the research program set forth in Chomsky (1989, 1992).

1. Previous analyses

Two analyses have been proposed for free relatives. The first analysis, which can be called the Adjunct analysis, was most clearly defended in Bresnan and Grimshaw (1978), and has been taken up again by Larson (1987). The Adjunct analysis of FRs claims that free relatives consist of a Wh- element with a CP adjunct as in (1b). The second analysis, which can be called the Null Head analysis, claims that the Wh- element of an FR is in Specifier position of a CP adjoined to a null head as in (1c). This analysis, first suggested by Kuroda (1968), was most successfully argued for by Groos & Van Riemsdijk (1978) and Hirschbühler (1978).

(1) a. I'll eat [what(ever) you order]2

b. 

c. 

[Diagram]

I would like to thank Richard Larson, Dave Lebeaux, Pierre Pica, Guido Vanden Wyngaerd, and Laurie Zaring for useful comments and discussions. The usual disclaimers apply.

1 The Adjunct analysis is known in the literature as the Head analysis. Jacobson (1990) calls this the NP-S analysis. The term Head analysis is potentially confusing since it wrongly suggests that an NP is functioning as an N. The Null Head analysis is known as the Comp analysis, a term equally confusing since updated versions of both analyses presumably need to refer to a CP projection for the sentential complement involved in FRs.

2 I will make no distinction between definite and universal (what(ever) interpretations of FRs. Jacobson (1990:4-5) shows that a sharp distinction between both interpretations would be an oversimplification. She shows that FRs introduced by whatever can have a definite meaning since these FRs can be anaphorically referred to by it, unlike universally quantified NPs such as everything. On the other hand, Wh- FRs without ever can have a universal meaning as in Do what the babysitter tells you to do (Jacobson 1990).
There are several theoretical and empirical arguments supporting the Null Head analysis over the Adjunct analysis that I cannot go through here (cf. Groos & Van Riemsdijk 1978, Hirschbühler 1978). Jacobson (1990) offers some recent empirical arguments against the Adjunct analysis. She points out that the Adjunct analysis does not explain why ordinary NPs cannot be followed by a CP adjunct. If a CP can be adjoined to the NP what, there is no reason it could not be adjoined to any other NP.

(2) a. *I'll order the food/ anything is recommended by the chef
b. I'll order what is recommended by the chef

Jacobson (1990) also notes that the Wh- element necessarily occurs with an associated sentential complement:

(3) a. I will read whatever/ all books
b. *Whatever/ All books are lousy

It is usually assumed that adjuncts are optional. If the CP is an adjunct to the Wh- NP, then why is it obligatory? it might of course be argued that the NP (what)ever licenses its CP adjunct in some special way, but it is hard to see how this can be achieved without stipulation. Under a Null Head analysis, this problem does not arise: whatever in (3) cannot modify No without an associated sentential complement because the Wh- element and the sentential complement necessarily form a single relative clause constituent.

Another argument comes from CP extraposition. If the clause following the Wh- NP were an Adjunct CP, we would expect it to be subject to extraposition in the same way as the CP complement of the subject NP in (4a):

(4) a. [Any reports t_j] will be published [that are on my desk by tomorrow];
b. * [Whichever reports t_j] will be published [that are on my desk by tomorrow];

However, (4b) shows that FRs do not allow for this type of extraposition.3

A more theory-internal argument is based on Rizzi (1990b). Following May (1985), Rizzi (1990b:378) assumes that the occurrence and position of Wh- elements at LF is determined by principle (5), the Wh- criterion (=Rizzi 1990b:(9)):

(5) a. Each +Wh- X° must be in a Spec-Head relation with a Wh- phrase
b. Each Wh- phrase must be in a Spec-Head relation with a +Wh- X°

This licensing of Wh- elements is a strong morphological requirement, in line with Chomsky's (1992) minimalist assumptions for motivating movement. If Rizzi's (1990b) Wh- criterion is taken seriously, the Wh- element of an FR also must be in a Spec-Head relation licensing the morphological Wh- property. Only the Null Head analysis is compatible with this idea: the Wh- NP necessarily is in the SpecCP position of the relative clause, where it is licensed by a +Wh- C°. Under the Adjunct analysis, it would have to be stipulated that the Wh- property of the Wh- NP is licensed in some other way. Ultimately, such a stipulation would boil down to an ad hoc distinction between two types of Wh- elements. Bresnan & Grimshaw (1978) and Larson (1987) effectively analyze elements such as whatever as ordinary universal quantifiers. Jacobson (1990:28) argues that despite their apparent semantic similarity to (universal) free choice any, elements such as whatever are very different from ordinary universals: they cannot be modified by almost or nearly, and they do not trigger negative polarity items such as ever. Moreover, Jacobson (1990) shows that the quantificational properties of FRs and interrogatives can be semantically analyzed in the same way. In the best of possible grammars, we would want to express a one-on-one relationship between morphological Wh- properties and semantic Wh- properties, rather than stipulate that the set of Wh- elements which must be licensed by a +Wh- C° in interrogative clauses can also independently function as universal quantifiers in FRs without needing to license their Wh- properties.

Despite these empirical and theoretical reasons for favoring a Null Head analysis, there are some important extraction facts which are incompatible with the Null Head analysis. Extraction of arguments out of FRs is generally impossible:

3 Thanks to Laurie Zaring for pointing this out to me.
(6) a. I will eat whatever the chef recommends to that person
   b. *This is the person to whom I will eat whatever the chef recommends
(7) a. I'd like to see whoever the chef recommends the paté to
   b. *This is the paté that I'd like to see whoever the chef recommends to

For the Adjunct analysis, this extraction is impossible because of the fact that an adjunct CP is a strong island for extraction (Cinque 1991). For the Null Head analysis, the impossibility of extraction is predicted by the Complex NP Constraint, hence by Subjacency (Chomsky 1986a).

However, there is a type of extraction out of FRs that is not predicted by the Null Head analysis. It seems that extraction out of FRs is possible out of the Wh- NP in a FR.4

(8) a. This is the author [of whom] I buy [[whatever books tj] the NYT recommends to its readers]
   b. This is the accident [of which] I read [[whatever reports tj] were published in the papers]

Under the Null Head analysis, this type of extraction should be just as impossible as the extraction in (6-7). If the Wh- NP in (8) is in SpecCP of a CP complement to a null head, the CP should be a Barrier to extraction out of the Wh- NP in (8) in the same way as in other relative clauses where extraction out of a Wh- NP in SpecCP is disallowed:

(9) a. I met an activist whose report of the nuclear accident
   the authorities ignored for weeks
   b. *This is [the nuclear accident] of which I met an activist
      [CP [whose report tj] the authorities ignored for weeks]

In other words, if FRs involved a null nominal head, the sentences (6b), (7b), (8ab) and (9b) should be all equally ungrammatical by the CNPC. The sentences in (8ab) thus also form a minimal pair with (10b), where ungrammatical extraction involves a constituent lower than SpecCP. Compare also with the much milder Wh- island violation in (10c):

(10) a. I buy whatever books the NYT recommends to its readers
   b. *These are the readers to whom I buy whatever books
      the NYT recommends
   c. ?? These are the readers to whom I know what books the NYT recommends

Summarizing, if FRs are headed by a null N°, the CP should be a Barrier to all extraction, a strong island in the sense of Cinque (1991). This is clearly not the case: extraction out of the Wh- NP is allowed and contrasts with the strong impossibility of extraction of complements out of the clause following the Wh- NP.

The Adjunct analysis of FRs of course predicts these extraction facts. Extraction out of the Wh- NP is possible because of the fact that the NP is governed and L-marked (Chomsky 1986a) by the matrix verb. Extraction out of the CP following the Wh- NP is impossible since the CP is an adjunct island not L-marked by the matrix verb (a strong island in the sense of Cinque 1991).

We are thus faced with a paradox: the extraction facts argue in favor of an Adjunct analysis of FRs, whereas the empirical and theoretical considerations noted above suggest that the Wh-element of an FR should be in SpecCP. How can we maintain the advantages of the Null Head analysis while at the same time accounting for the extraction facts?

4 It has been pointed out to me that the following sentence, where the preposition has been stranded in the Wh- NP, is entirely ungrammatical:
   i. *This is the author who I buy whatever books of the NYT recommends to its readers
   However, I think this ungrammaticality must be explained independently of the FR construction. The same ungrammaticality shows up in indirect interrogatives:
   ii. *This is the author who I know whatever books of the NYT recommends to its readers
   iii. This is the author of whom I know whatever books the NYT recommends to its readers
2. FRs and indirect Wh- clauses are structurally identical

In order to solve the structural paradox posed by FRs, I would like to claim that FRs are CPs without nominal heads. This proposal amounts to saying that FRs are structurally identical to indirect Wh- clauses, i.e. interrogatives and declaratives as in (11) (cf. Bresnan 1972):

(11) a. Murasaki wondered what Genji had written/ how she played the lute
    b. Murasaki told me what Genji had written/ how she played the lute

In this section, I would like to focus on the motivation for a bare Wh- CP analysis of FRs and indirect Wh- clauses, showing that this proposal preserves the advantages of the Null Head analysis and the Adjunct analysis.

Of course there are a number of important differences between indirect Wh- clauses and FRs which immediately come to mind. Before focusing on these important differences in section 2.2., I would first like to point out the striking similarities between FRs and indirect Wh- clauses. These similarities involve extraction facts, morphological properties and quantificational properties. The structural identity between FRs and indirect Wh- clauses will be taken to account for these similarities. In section 2.3, it will be shown that the important differences between indirect Wh- clauses and FRs, which at first sight prevent their analysis as bare Wh- CPs, can be elegantly resolved in a framework making use of Generalized Transformations in the sense of Chomsky (1992).

2.1. Similarities between FRs and indirect Wh- clauses

First of all, extraction out of a Wh- NP in SpecCP of an indirect Wh- clause is allowed. Chomsky (1986a:26) states that a matrix verb must be allowed to L-mark the specifier in a structure such as (12) in order to explain sentences such as (13):

(12) V [CP Wh- phrase C EP] (=Chomsky 1986a:(50))

(13) a. Este es el autor [del que]i no sabemos [CP [qué libros ti ] leer]
    'This is the author by whom we don't know what books to read'
    (=Chomsky 1986a:(48a), citing Torrego 1985)
    b. ¿De qué autora no sabes qué traducciones han ganado premios internacionales?
    'By which author don't you know what translations have won international awards?' (=Chomsky 1986a:(49b), citing Torrego 1985)

Chomsky (1986a) states that if the verb saber ‘know’ in (13) does not L-mark the Wh- element in SpecCP, the sentences should be ruled out by subjacency, since the Wh- element in SpecCP, and by inheritance CP itself, would then be Barriers to movement. It is important to point out that Spanish FRs allow extraction out of the Wh- NP in the same way indirect Wh- clauses do:

5 The restriction to the singular in the FR cuanta traducción ha ganado... has to do with a more general restriction in Spanish: universal quantification always requires the singular, e.g. ninguna cosa ‘nothing’ vs. * ninagas cosas. The question arises as to why a similar construction is impossible with French quel(le)(s) (or, for that matter, with Spanish cuales):

i. * Voilà l'auteur dont j'ai lu quelles traductions ont gagné des prix internationaux
   'This is the author of whom I read those translations (that) have won international prizes'

However, similar sentences without extraction are also excluded:

ii. * J'ai lu quelles traductions/ lesquelles ont gagné des prix internationaux
   'I read those translations/ those (that) have won international prizes'

Quel(le)(s) ‘what’ only shows up in indirect Wh- clauses:

iii. Je me demande/ lui ai dit queles livres tu as lu
    'I wonder/ told him which books you have read'

The most likely explanation for this difference between FRs and indirect Wh- clauses in French is that quel(le)(s) ‘what’ is incompatible with the definite/ universal quantification of FRs. French does not productively have a morpheme attaching to Wh- elements such as English ever. In French, only qui ‘who’ can spell out the universal property in micromorph ‘whoever’, where quiue is diachronically a temporal adverb (Bouchard & Hirschbühler 1986).
Generalized Transformations and the Wh-cycle

(14) a. ¿De qué autora has leído cuanta traducción ha ganado premios internacionales?
   'By which author did you read whatever translation won international awards?'
b. Of which author do you read whatever publications you can find?

The possibility of extraction out of the Wh-NP in FRs was also discussed in the previous section. The examples in (14) are parallel to those in (8).

Similar movement out of a Wh-NP in SpecCP is relevant for Binding. Chomsky (1992:54) claims that LF movement of self (LF cliticization or CLt) out of the Wh-NP accounts for the fact that the anaphor can be bound by the matrix subject in (15):

(15) John wondered [which pictures of himself] Bill saw t (=Chomsky 1992:(36))

Notice that the same Binding facts are attested in FRs, suggesting that the same analysis in terms of LF extraction applies:

(16) John gave me whatever pictures of himself Bill had found

The Binding facts in indirect Wh-clauses and FRs thus seem to mirror the extraction facts. Both overt and LF extraction is possible out of the Wh-NP of FRs and indirect Wh-clauses. If the Wh-NP of both FRs and indirect Wh-clauses is in a SpecCP which is directly governed by the matrix verb, extraction out of the Wh-NP in (8-13-14) and (15-16) can be straightforwardly accounted for along the lines suggested by Chomsky (1986a, 1992).

Secondly, it has often been noted that many languages have identical Wh-morphemes in the NP domain for FRs and indirect Wh-clauses. In French, the referential properties of the set of Wh-elements replacing NPs is identical for questions and FRs:

(17) a. Je me demande [qui C° tu as vu ] 'I wonder who you saw'
b. Je me demande [ ce que tu as vu ]7 'I wonder what you saw'

(18) J'ai vu [qui/ ce que tu as vu ] 'I saw who/what you saw'

(19) a. L'homme/ le train [ O qui est arrivé]
   'The man/ the train who/ which arrived'
b. L'homme/ le train [ O que tu as vu]
   'The man/ the train that you saw'

Importantly, FRs do not use an empty operator strategy as do relative Wh-clauses.8 The Wh-words in (17-18) behave as pronouns with their own reference (+animate qui 'who' -animate ce que 'what'). Relative que/qui are simply complementizers which do not express an animacy distinction, with que changing to qui when a subject is relativized (Kayne 1976, Rizzi 1990a).

This correspondence between FRs and indirect Wh-clauses is not predicted by either the Adjunct analysis or the Null Head analysis. As I have noted above, under the Adjunct analysis, it is simply a coincidence that the set of Wh-words, which normally should be related to a +Wh-C°, can also independently function as NPs. Under the Null Head analysis, the correlation between FRs and indirect Wh-clauses is even stranger. Why would a relative clause to a null N° head require overt Wh-NPs in SpecCP rather than the null operators of other relative clauses?

6 With the exception of quiconque 'whoever', which can only be used in FRs. This is due to the universal quantificational force of quiconque 'whoever'. Also, pourquoi 'why' and comment 'how (instrument)' can appear in indirect Wh-clauses, but not in FRs. For an explanation of this restriction, see Larson's (1987) analysis of the nonexistence of English FRs with why (cf. footnote 10)

7 Contra Bellier (1989) and Friedemann (1989), Pollock (1992) argues that ce in interrogative ce que clauses cannot be the nominal head for a que relative clause, since ce que clauses show weak Wh-island effects and not the stronger CNPC effects which would be expected if ce que clauses were complex NPs. Pollock (1992) suggests that ce is a D° selecting a complement clause (cf. also Zaring 1992). See Rooryck (1992) for arguments that ce actually is an AGR-C°, spelling out the case assigned to CP. No matter what the correct analysis of ce, it is clear that the complex ce que is used in both indirect Wh-clauses and FRs. Therefore, it is justified to view them as a complex Wh-NP at the descriptive level.

8 Prince (1989) shows that Yiddish question morphemes are also different from relative morphemes. FR Wh-morphemes are identical to question morphemes, not like relative pronouns. See Horvath & Grosu (1987) for similarities between FRs and indirect Wh-clauses in Rumanian.
Generaliized Transformations and the Wh- cycle

It might be argued that the null operator strategy cannot be used here because of the need for identification of the null N° heading the free relative. However, it is entirely unclear how an overt Wh- phrase in Spec, CP would be better able to identify or license the null N° than a null operator. Null operators are referentially 'strong' enough to trigger morphological changes on C° as attested by the que --qui alternation in French, which can be explained by Spec-Head agreement of C° with the operator in SpecCP and agreement of C° with AGR-S° (Rizzi 1990a:55-56). The complementizer bearing overt nominal subject agreement does not have o-features which are different from an overt Wh- NP originating in subject position. The only features in which overt Wh- NPs differ from empty operators concern referential features such as animacy, the Wh- NP qui 'who' being +animate and ce que 'what' being -animate. It is unlikely that these features would be necessary to license the empty N° head. In other cases where null NP categories are present, such licensing never appears to be necessary. In cases of control by empty NP arguments, for instance, the ± animacy of a null N° head is determined by the governing verb (Rizzi 1986):

(20) Ceci amène [N° e] à conclure les choses suivantes
    'This leads to conclude the following'

If FRs have the structure in (1c), one would expect animacy features of the empty N° to be also determined by the selectional restrictions of the matrix verb. These features of N° should be enough to license null operators in SpecCP of the FR clause: there is no reason to assume that an empty N° with selectionally determined features behaves in a way that would be substantially different from an overt N°. Under the assumption that licensing is a contentive relation, there seems to be no property of the empty N° heading FRs that could only be licensed by an overt Wh- NP. If we want to maintain that the overt Wh- NP is moved to SpecCP because of the need for identification of N°, it would have to be stipulated that the empty N° is not accessible for the matrix V° and can only be identified by an overt Wh- NP. Besides stipulatively 'closing off' the NP in which the FR is contained, this analysis would of course prevent any explanation of the Matching phenomenon (cf. infra). Under a Null Head analysis, then, the morphological correlation between the Wh- NPs appearing in FRs and indirect Wh-clauses is entirely unexpected. If both FRs and indirect Wh-clauses are structurally analyzed as bare Wh- CPs, the morphological identity of the Wh- elements involved in both Wh-constructions follows without stipulation.

A third way in which FRs and indirect Wh-clauses are similar involves their semantics. Adapting Cooper (1983), Jacobson (1990) suggests that FRs and Wh-questions have a similar meaning which should be distinguished from the meaning of a relative clause. Jacobson (1990:15) suggests that a relative clause such as which I ate denotes the set of individuals which I ate. Relative which therefore is an identity function on properties (Jacobson 1990:16). The Wh- constituent what John ate however has as its predicative meaning the set of maximal plural entities that John ate. This predicative expression then shifts into an NP denoting the maximal plural entity that John ate, allowing Jacobson (1990) to derive the fact that FRs can be sometimes definite (I'll order what you are eating) and sometimes universal-like (I'll order whatever you eat). Jacobson (1990) extends this analysis of FRs to Wh-questions (and presumably indirect Wh-clauses in general). She claims that a Wh-question basically has the same meaning as the one proposed for unshifted FRs, but in this case the predicative Wh-expression semantically shifts to a proposition p such that there exists some entity X such that p is true, and the denotation of the Wh- constituent is true of X. Jacobson goes on to show that FRs and indirect Wh-clauses share a number of other semantic properties illustrating that a semantically unified analysis of both is desirable. What Jacobson (1990) shows is that the Wh-elements in FRs and indirect Wh-clauses have quantificational properties which are absent in the case of relative clauses. This raises the question as to whether the syntactic analysis of FRs and indirect Wh-clauses should not to some extent reflect this semantic unification. Notice that such a semantically unified analysis of FRs and indirect Wh-clauses is totally unexpected.

9 This view of Wh-elements in questions and FRs as having universal quantificational force has been challenged by Berman (1989). Interestingly, however, Berman (1989) analyzes the semantics of indirect questions and FRs together in terms of a notion of quantificational variability. See Lahiri (1990) for a refinement of Berman's (1989) position.
Generalized Transformations and the Wh-cycle

Generalized Transformations and the Wh-cycle

on a Null Head analysis. The Null Head analysis predicts a strong correlation between relative clauses and FRs, contrary to fact.

In the same vein, Larson (1987) points out that the Null Head analysis does not account for the quantificational properties of FRs. Larson (1987:263), who adopts the Adjunct analysis for FRs, notes that the status of the Wh-element as a quantifier is an 'absolute' property that should be independent of the predicate exercising selectional restrictions on this Wh-element over and above the null N° head (Groos & Van Riemsdijk's (1978) Comp accessibility). Proponents of the Null head analysis could of course claim that the null head determines the quantificational properties of the Wh-element in FRs, but it is hard to see how this could be achieved without stipulating the existence of null N° quantifiers, both definite and universal-like. If FRs and indirect Wh-clauses are bare Wh-CPs, no such stipulation is necessary, and the semantic similarity of Wh-elements in both constructions with respect to quantification is in line with their identical syntactic representation as indirect CPs.

2.2. Differences between FRs and indirect Wh-clauses

Let us now turn to the differences between FRs and indirect Wh-clauses which seem to preclude an analysis claiming structural identity for both. These differences involve extraction and the Matching effect. First of all, despite the fact that extraction out of the Wh-NP in both FRs and indirect Wh-clauses is possible, extraction of arguments out of the clause following the Wh-NP gives rise to strongly ungrammatical sentences in the case of FRs (10b), while the same type of extraction only yields much weaker Wh-island violations in the case of indirect Wh-clauses as in (10c). As noted above, both the Adjunct analysis and the Null Head analysis of FRs immediately derive this crucial difference between FRs and indirect Wh-clauses, since in both cases the CP involved in the FR is claimed to be a strong island in the sense of Cinque (1991). By equating the structure of FRs with that of indirect Wh-clauses, the analysis proposed here predicts that extraction out of FRs should only give rise to weak Wh-island violations, contrary to fact.

A second problem for analyzing FRs as bare Wh-CPs lies in the Matching phenomenon: unlike indirect interrogatives, free relatives require that the phrase introducing the relative clause conform to the selectional restrictions and subcategorization requirements of the governing verb. This can be illustrated by (21) which obeys matching, and (22) which does not:

(21)  a. I called who you asked me to talk to
       b. J'ai rencontré qui tu voulais que je rencontre
           'I met who you wanted me to meet'

(22)  a. *I called what you asked me to do
       b. *J'ai rencontré ce que tu m'as dit de faire
           'I met what you told me to do'

In some languages, Matching also requires the case of the Wh-element to be identical to that of the position the FR occupies. Matching does not apply in interrogative clauses, where any Wh-element may appear in Spec-CP position.

(23)  a. I told him something/*someone
       b. I told him what you are doing/*who you met/*how you met

Indirect Wh-clauses are selected as complements of a specific set of verbs expressing (know), presupposing (say, tell) or entailing knowledge (wonder, inquire, ask), which assign a +Wh-value to C°. This is clearly not the case for FRs: FRs can appear in any position, as subjects complements or adjuncts. Moreover, Larson (1987) convincingly shows that FRs in English reduce to the categories NP and AP, claiming that in structures such as in whatever town you live, there is no Wh-PP in whatever town, but rather a Wh-NP selected by the preposition

---

10 The declarative or interrogative value of the Wh-clause is determined by the particular verb selecting the Wh-clause (Baker 1970, Bresnan 1972). The Wh-feature of C° should be dissociated from the interrogative interpretation.
This analysis certainly does not apply to indirect Wh- clauses as in (24), where there clearly is a Wh- PP in SpecCP.

(24) I wondered in what town you gave a talk

If Larson (1987) is correct, FRs can be generated in any subject, complement or adjunct position that is also an NP or an AP position. Moreover, the Wh- NP/ AP has to satisfy the selectional restrictions of the position the FR is in. Indirect Wh- clauses are restricted to the complement positions of a specific set of verbs and allow for any Wh- element to be in SpecCP position. If both FRs and indirect Wh- clauses are bare Wh- CPs, an explanation for this contrast is needed.

2.3. Explaining the differences: GT and Wh- CPs

I would like to claim that the difference in extraction properties and distribution of FRs and indirect Wh- clauses can be accounted for in the minimalist framework of Chomsky (1992). Chomsky (1992) reintroduces the powerful mechanism of Generalized Transformations (GT). GT is a substitution operation which targets a phrase-marker K, adds $\emptyset$ to form $K^\circ$ which must conform to X' theory. In this way, an X' can be extended to an XP, by adding a specifier position, or an $X^\circ$ can be extended to an X' adding a complement position. The $\emptyset$ in this position can then be replaced by another phrase-marker by substitution, or, in the case of specifier positions, by moving an XP to substitute $\emptyset$.

A question immediately arises as to the nature of the phrase-markers which can be inserted by GT to replace $\emptyset$. What phrase-markers does GT have access to? In case $\emptyset$ is replaced by a CP, for instance, the question arises as to whether the CP phrase-marker has already been internally subject to movement processes or not. Nothing in the definition of GT prevents GT from having access to both CPs in which movement has not yet taken place, and CPs in which movement has already taken place before insertion. In other words, in principle, GTs must be allowed to insert both 'finished' XPs (in which both overt and LF movement has already taken place) and 'unfinished' XPs (in which movement still has to apply). I would like to argue that GTs can insert both types of phrase-markers. In the context of FRs and indirect Wh- clauses, I claim that the syntactic differences between FRs and indirect Wh- clauses can be explained by exactly this distinction: whether or not movement has applied in the CP before or after GT insertion.

Free relatives are CPs in which Wh- movement has already taken place before insertion. Note that the notion of the strict cycle does not prevent this. In these CPs, SpecCP is filled by a Wh-XP which shares its features by Spec-Head agreement with $C^\circ$. I assume with Rizzi (1990a) that there is an AGR element in $C^\circ$, perhaps an AGR-$C^\circ$ projection (Rivero 1988). As a consequence, the features of $C^\circ$, which are those of the element in SpecCP must match the c-selectional features of whatever position the CP occupies. Since the element in SpecCP is accessible for government, s-selectional restrictions can be imposed upon it. Both selections account for the matching effect in (21-22). For the verb selecting the FR, the presence of a Wh- NP in SpecCP of the FR makes the entire CP projection function as if it were an NP categorically. Since when is a Wh- element of the category NP (Larson 1987), a FR whenever clause can take the place of a temporal adjunct that is of category NP. In English and French, the only XPs which will be able to fully agree with AGR-$C^\circ$ are NP and AP.12 This is why FRs can be inserted in any subject, complement, or adjunct position that is also an NP position. PPs in SpecCP of FRs would not be able to agree in nominal features with AGR-$C^\circ$.

11 This allows Larson (1987) to explain why there are no FRs with why: he shows that contrary to when and where, why is not a bare NP adverb. The same analysis applies to French pourquoi 'why' and comment 'how (instrument)', which can also appear in indirect Wh- clauses, but not in FRs. Notice that Larson's (1987) insight that there are no PP FRs can be easily implemented in our analysis of FRs as indirect Wh- clauses: it suffices to say that the indirect Wh-clause whatever town you live is selected by the preposition in. Note also that Larson (1987) does not explain why FRs are limited to NP and AP.

12 See Chomsky & Lasnik (1991) for arguments to maintain the distinction between c-(categorial) selection and s-(semantic) selection. See Grimshaw (1979), Pesetsky (1982), and Chomsky (1986h) for arguments in favor of deriving c-selection from s-selection, and Emonds (1992) for arguments to the contrary.
preventing PPs from showing up in this position. This is what explains Larson's (1987) restriction of FRs to NPs and APs.

How are indirect Wh- clauses derived? Indirect Wh- clauses are CPs which are inserted before movement applies to them. This means that when the CP is inserted, the verb only governs C°, SpecCP being empty. The verb will therefore impose selectional restrictions on C° only (+ Wh-; interrogative, declarative, etc). Any + Wh- argument or adjunct XP will be able to move to SpecCP in order to check + Wh- properties. Since selectional properties cannot be changed in the course of the derivation, the Wh- element in SpecCP will not be subject to selectional restrictions of the governing verb: once + Wh- selection of C° has taken place, it remains the same throughout the derivation. Therefore, any Wh- element can show up in SpecCP of an indirect Wh- clause, without being subject to selection under government from the matrix verb.

The contrast between FRs and indirect Wh- clauses with respect to extraction facts still has to be explained. Recall that overt and implicit extraction out of the Wh- NP in SpecCP is allowed in both indirect Wh- clauses and FRs (8-14-16). Extraction of constituents out of the clause following the Wh- NP gives rise to strong Subjacency violations in the case of FRs (10b), whereas similar extractions out of indirect Wh- clauses only give rise to weak Wh- island violations (10c). This situation cannot be explained by the strict cycle condition: in principle, extraction from the lower CP to the higher domain should always be possible on the second cycle, whether a GT inserted CP has internally undergone movement or not. The analysis advocated here needs to explain why extraction of constituents out of the clause following the Wh- element is so much worse for FRs than for indirect Wh- clauses.

I have proposed that a FR has been subject to all syntactic movement operations before GT insertion. In other words, GT has access to 'syntactically finished' constituents just before spell-out, and can reintroduce these 'finished' constituents into phrase-markers. I would like to suggest that CPs in which all movement has taken place before insertion become opaque to further syntactic operations.

It is natural to assume that a syntactic constituent which has internally been subject to all syntactic operations cannot be again subjected to these operations when reinserted into the syntax. This requirement can be viewed as an economy principle in the sense of Chomsky (1989). This principle can be formulated as in (25):

(25) The Double Jeopardy Principle (DJP)

Constituents which have internally been subject to all syntactic operations cannot again be subjected to them when reinserted by GT.

Once a 'finished' Wh- CP (a FR) is inserted by GT, the DJP ensures that no elements can be extracted from it, since this CP is opaque to syntactic operations. However, there is one position in FRs that is still accessible for extraction, namely SpecCP. Chomsky (1986a) assumes that a Wh- element in SpecCP is properly governed and L-marked by the verb selecting that CP. Therefore, extraction from the Wh- element in SpecCP is allowed. The same is true when the FR is governed by a verb that L-marks it (cf. (8), (14-16)). Proper government therefore partly overrides the 'fixed' character of FRs: this is only what is expected, since the verb must have a way of exercising selectional restrictions on the FR in order to derive Matching phenomena. The C° of FRs therefore is properly governed and the Wh- element in its SpecCP is L-marked. As a consequence, the Wh- NP in SpecCP of the FR is accessible for extraction as in (14-16). However, anything that is further down in the FR is invisible for extraction by virtue of (25). The DJP overrides the L-marking of the entire free relative CP, while government of the verb into SpecCP overrides the DJP. This asymmetry will be explained below. It is important to stress that under this approach, the strong impossibility of extraction of constituents out of the clause following the Wh-element in FRs does not follow from Subjacency, but from (25). The DJP does not apply to indirect Wh-clauses, since these are inserted as CPs in which all syntactic processes apply after GT insertion, and as such the syntax will apply only once to them.

The DJP also accounts for a problem raised by the bare CP analysis for FRs that has not yet been addressed. If FRs are bare Wh- CPs, the Wh- element in SpecCP should in principle be able to move on to the SpecCP of the matrix clause by movement on the second cycle. This is not the case:
(26) a. I told you that I will eat [CP what [IP you order t₁]]
   b. * I told you what I will eat [CP t₁ [IP you order t₁]]

Note that the trace in SpecCP is antecedent governed by the Wh-element in the higher SpecCP. The problem therefore does not lie with the movement on the second cycle itself. Let us now assume with Chomsky (1992:21) that the basic syntactic operation is not movement but Form-Chain, yielding the chain (What₁ - t₁' - t₁) in (26b). By virtue of the DJP, this chain is illegitimate under our analysis of FRs as GT inserted Wh-CPs. The first part of the chain (What₁ - t₁') is formed on the second cycle and is perfectly all right. However, the second part of the chain (t₁' - t₁) involves a double application of Form-Chain: Form-Chain has applied once before GT insertion inside the FR yielding the chain (What₁ - t₁) in (26a), and would have to apply a second time after GT insertion to the (t₁' - t₁) subpart of the chain (What₁ - t₁' - t₁). This second application is prohibited by the DJP, since the chain in the FR cannot again be subjected to Form-Chain. Notice that in a case of successive cyclic movement in Wh-clauses such as *Who did you wonder whether Mary saw t₁ there is only a single application of Form-Chain: the Wh-clause is GT-inserted before any movement takes place, and movement (Form-Chain) applies in a single application of the syntax, proceeding from the first to the second cycle. It follows then that the only type of movement permitted out of an FR must originate in the Wh-element itself as in (14-16), since this is the only chain that can 'reach into' the FR without violating the DJP. In the case of (14-16), the chain relating the extracted element and its trace in the Wh-NP in SpecCP of the FR is entirely formed on the second cycle. The constraints on extraction out of FRs illustrate how the DJP functions as an economy principle forcing the syntax to apply one time only to syntactic strings.

At first sight, the DJP might seem stipulative. However, there seem to be other syntactic environments in which the DJP can be argued to apply. The first case involves Aux-to-Comp cases in Italian. Italian Aux-to-Comp constructions seem to constitute strong islands in the sense of Cinque (1991); the extraction in (27b) does not have the flavor of a typical (weak) Wh-island violation, but is much stronger. (28ab) constitute a minimal pair:

(27) a. Ritengo [cp aver Lia risolto molti problemi]
   'I consider have Lia solved many problems'
   b. * Questi sono i problemi che ritengo aver Lia risolti
   'These are the problems that I consider have Lia solved'

(28) a. ??Questo e il ragazzo a cui mi domando come si possa regalare dei fiori
   'This is the boy to whom I wonder how one can give flowers'
   b. * Questo e il ragazzo a cui ritengo aver Lia regalato dei fiori
   'This is the boy to whom I consider Lia to have given flowers'

The strong islandhood of Aux-to-comp constructions is entirely unexpected. Small clause complements of verbs such as ritenere 'consider' clearly are theta-marked and L-marked, since they allow for extraction of the small clause subject:

(29) a. Ritengo [sc Lia intelligente]
   'I consider Lia intelligent'
   b. Questa e la persona che ritengo [sc t intelligente]
   'This is the person that I consider intelligent'

If the SC in (29) is theta-marked and L-marked, there is every reason to believe that the Aux-to-Comp CP in (27) is also theta-marked and L-marked. Why would Aux-to-Comp constructions not allow for successive cyclic movement of arguments into the matrix clause despite their being theta-marked and L-marked? In the spirit of the proposal investigated here for FRs, I would like to propose that Aux-to-Comp CPs are CPs which have been inserted after the syntax has applied inside the CP. Aux-to-Comp in the CP then takes place before GT-insertion of the CP as a complement of ritenere 'consider'. The DJP will prevent any movement out of the Aux-to-Comp CP, since the syntax cannot apply a second time to the Aux-to-Comp CP. As a result, the Aux-to-Comp CP behaves like a strong island despite its being theta-marked and L-marked. The unexpected strong island in Aux-to-Comp constructions such as (27-28b) can be taken to independently support the existence of the DJP.
A similar argument for the DJP comes from Instrumental adjuncts. Baker (1988:243) has shown that Instrumental adjuncts have argument-like properties with respect to the ECP in that they can escape Wh- islands:

(30) a. (?) With which key do you always forget how to open doors?
    (Baker 1988:35b))

b. (?) This is the key without which I don’t know how to open the door.

This means that they must be theta-governed by the verb, a likely assumption, since Instruments have a thematic relation with the verb selecting them (Baker 1988). Being theta-governed, Instruments also must be L-marked (Chomsky 1986a). Now, the L-marking of Instruments predicts that it should be possible to extract out of them. However, this prediction is not borne out:

(31) a. * This is the person to whom I helped my sister by giving some money
t
b. * To whom did you leave without speaking t (=Cinque 1991:1(3a))

In Cinque’s (1991) terms, adjuncts are always ‘strong’ islands. In view of the contrast between (30) and (31), a Barriers-type framework has to stipulate that adjuncts are not L-marked despite their being theta-governed. The minimalist framework does not require such a stipulation, provided a principle such as the DJP is accepted. Lebeaux (1988) has suggested that adjuncts are inserted by GT (Chomsky 1992). If it is assumed that all movement in the adjunct has taken place before its GT insertion, the impossibility of extraction out of the adjunct after GT insertion can be derived in the same way as in the case of FRs. In both cases, extraction would involve applying a syntactic operation to ‘syntactically finished’ constituents which have already been subject to all syntactic processes before GT insertion. It seems then that the DJP formulated in (25) can be justified on sufficiently independent grounds. The DJP can be viewed as an economy principle preventing recursive application of the syntax after GT reinsertion.

3. Some residual problems: nonmatching; infinitival FRs

Finally, I would like to discuss how the difference between matching and nonmatching languages can be implemented in the analysis of FRs presented here. In nonmatching languages, the SpecCP of FRs can be occupied by Wh- elements with a case that does not correspond to the case corresponding to the position of the FR. Similarly, in such languages PPs can occur in the SpecCP position of an FR. Harbert (1983a) quotes the following examples from Gothic:

(32) a. ushafjands [NP [PP ana þammei] lag] (Luk. 5:25)
    picking up on which he-lay
    ‘Picking up that on which he lay’ (=Harbert 1983a:(13))

b. þan-ei (=sa þan-ei) frijos siuks ist (Joh. 11:3)
    Acc-compl (Nom Acc) you-love sick is
    ‘(The one) whom you love is sick’ (=Harbert 1983a:(18a))

In (32a), a PP occurs in the SpecCP position of an FR which is in object position, and in (32b), the case of the NP in subject position is accusative, corresponding to the case of the object gap in the FR rather than to the nominative case required for subjects.

The analysis of FRs proposed in the preceding sections only applies to matching languages. The difference between matching and nonmatching languages has been a source of considerable debate (Groos & Van Riemsdijk 1979, Bresnan & Grimshaw 1978, Harbert 1983ab, Hirschbühler 1978, Hirschbühler & Rivero 1981, 1983, Sufier 1984, Horvath & Grosu 1987, Grosu 1987). It is clear then that something more needs to be said to accommodate nonmatching languages in the bare Wh- CP analysis adopted here for FRs. Harbert (1983ab) has noted that nonmatching languages are usually languages with a rich case system (Russian, Classical Greek, Gothic). Adopting the Null Head analysis, Harbert (1983ab) therefore suggests that in nonmatching languages the null nominal head of FRs, a pro, can be identified by a Wh- element with ‘rich’ Kase (where Kase includes both case-marking and Ps), in
conjunction with person and number agreement. Grosu (1987) has pointed out that Finnish, which has a rich case-system, only has restricted nonmatching. Grosu (1987) essentially refines Harbert's notion of case-identification by suggesting that in nonmatching languages identification involves person and number features, but not Kase, whereas in Matching languages identification affects Kase, but not person and number features. Grosu (1987) moreover shows that in Finnish and Rumanian exhibit Kase hierarchies in nonmatching FRs that can be accounted for in the system he adopts.

Note that Matching phenomena have been analyzed in § 2.3. as a result of Spec-Head agreement in the domain of AGR-C-P between the nominal features of the Wh- NP and a nominal AGR-C°. Harbert's (1983) and Grosu's (1987) analysis for nonmatching in which Kase identification takes place via predication between a null nominal pro head and a Wh-element in SpecCP of a relative clause can be replaced without cost by an analysis in with Kase identification takes place by Spec-Head agreement between a Wh-element in SpecCP and AGR-C°. In the analysis presented here, AGR-C° simply plays the role of Harbert's and Grosu's pro. The agreement interaction between a Kase-marked Wh- NP in SpecCP and AGR-C° then can be subject to parameterized variation in the same way Harbert's (1983) and Grosu's (1987) Kase identification is. In our analysis, the difference between matching and nonmatching is a function of Spec-Head agreement in the AGR-C-P projection. Granted a sufficiently rich variety of Spec-Head agreement phenomena in the domain of AGR-C-P, Harbert's (1983) and Grosu's (1987) analyses in terms of case-identification can simply be recast in a framework assuming the analysis of FRs as bare Wh- CPs. For all practical purposes, the Harbert/ Grosu analysis and the one presented here are technically equivalent with respect to (non)matching phenomena.

A last difference between FRs and indirect Wh-clauses that should be discussed here concerns their tensed or untensed character. Pesetsky (1982) analyzes infinitival FRs in Russian as bare Wh-CPs which are subject to a rule of QR satisfying selection at LF. Following Pesetsky (1982), Grosu (1987:52-54) offers evidence from Spanish and Rumanian that infinitival and subjunctive FRs are not FRs at all, but indirect Wh-clauses, contra Suner (1984). (33) a. Andrea no tiene [con quién salir (INF)] (=Grosu 1987:(20a))
   'Andrea does not have (anyone) with whom to go out'
  b. Andrea nu are [cu cine vota (INF)] (=Grosu 1987:(20bc))
  c. Andrea nu are [cu cine să voteze(SUBJ)]
   'Andrea does not have (anyone) with whom to vote'

Importantly, Grosu (1987) observes that these so-called infinitival FRs have indefinite meaning, are nonmatching, and are restricted to a small number of verbs. Therefore, I will continue to assume with Grosu (1987) that there are no true infinitival FRs. At first sight, this establishes another difference between FRs, which can only be tensed, and indirect Wh-clauses which can be both tensed and untensed. Note that this difference is as much a problem for the Null Head and the Adjunct analyses as it is for the analysis proposed here: since these analyses presuppose a relative clause analysis, they predict that infinitival relatives should be possible as either complements to a null head or adjuncts to a Wh-XP. However, it seems that the lack of tense is not a property that is exclusively related to FR Wh-clauses. There are at least some indirect Wh-clauses that do not have untensed counterparts (cf. Rooryck 1992):

(34) I really love (it) when you sing that song/ *when to sing that song

Notice that this indirect Wh-clause has a universal-like interpretation, in the same way FRs can have universal-like interpretations. The absence of infinitival FRs can probably be related to the quantificational properties FRs display. Rooryck (1992) notes that infinitival Wh-CPs have a deontic meaning, unlike their tensed counterparts: I know what to read means I know what I have to read, not I know what I will read. This deontic value of infinitival Wh-clauses can be viewed as a quantifier-like modality expressed in C°. It is not difficult to see how this deontic C° value of infinitival Wh-clauses would conflict with the universal-like value of the Wh-elements in free relatives and constructions like (34): the CP projection would have to accommodate two different universal quantifiers. The absence of infinitival FRs should be
derived from the interaction of the modal properties of Wh-infinitives with the universal-like properties of FRs.

4. Conclusion

Summarizing, I have shown that an analysis of both FRs and Wh-CPs as bare Wh-CPs can explain the strong similarities between both constructions with respect to extraction facts, quantificational characteristics, and morphological properties. The most important syntactic differences between FRs and indirect Wh-clauses, involving extraction and Matching, can be derived from the way in which both types of Wh-CPs are inserted in a sentence. Indirect Wh-clauses are Wh-CPs in which movement takes place after GT insertion. FRs are not relatives at all, but GT inserted Wh-CPs in which movement has taken place before GT insertion. It thus seems that the term 'free relative' itself really is a misnomer.

References

Hirschbühler, Paul. 1976. Two analyses of free relatives in French. Papers from the sixth meeting of the NELS, Montréal papers in Linguistics III, 137-152.


