Three Stages in the Derivation of Free Relatives

Anke Assmann[*]

Abstract

In this paper, I argue for a Comp account of free relatives that is based on the idea the covert external head of the relative clause is part of the *wh*-phrase. Thereby, the approach derives the close link between the covert head and the *wh*-phrase and is able to account for the puzzling behavior of free relatives with respect to the position of the *wh*-phrase. Concretely, I assume that the derivation of a free relative clause proceeds in three stages: in the first stage, the covert head is created out of the *wh*-item in the lexical array via pre-syntactic copying. In the second stage, the syntactic structure is built, whereby the *wh*-item is merged inside the embedded clause and the covert head is merged as the CP-external head of the relative clause. Finally in the third stage, the additional features created by pre-syntactic copying are deleted post-syntactically.

1. Introduction

Free relatives (FR) – together with parasitic gap and ATB constructions – belong to a class of phenomena where the position of certain items that have undergone movement is not clear. In FR constructions, the item in question is the *wh*-phrase, which appears to stand for the nominal head of the relative clause as well as the relative pronoun. Since it fulfills the function of both an item that is outside the embedded clause and an item that is inside the embedded clause, determining the position of this phrase presents an interesting task for syntacticians. Looking at various properties of FRs leads to a paradox though: the *wh*-phrase seems to be part of the embedded clause as well as the matrix clause.

An example of an FR in German is given in (1).[1]

(1)   Ich werde niemandem zeigen was   ich gefunden habe.
       I   will   nobody.DAT show what I   found   have
       'I won't show to anybody what I found.'   *(Ott 2011: 184)*

(1) shows the two most prominent properties of FRs that distinguish them from normal headed relative clauses: the lack of an external nominal head and the form of the relative pronoun, which appears in the form of a *wh*-phrase. FRs can be found in various languages. In what follows, the data I am concerned with are mainly drawn from German, but the observations made for them hold cross-linguistically.

The aim of this paper is to develop a new approach to FRs that is capable of deriving their puzzling properties, which seem to result from the ambivalent function of the *wh*-phrase. In order to analyze FRs, I will make use of a very traditional means of syntactic analysis: timing of operations. Concretely, I will assume that there are three stages in the derivation that roughly correspond to the traditional notions of the pre-cycle, the cycle, and the post-cycle (cf. McCawley 1970, Postal 1972, Pullum 1979): the lexical array (pre-cycle), the syntactic derivation (the cycle) and a post-syntactic component (post-cycle). Crucially, I will suggest that the copy operation, which applies in the syntax as part of the complex process of movement (Internal Merge), may also apply pre-syntactically in the lexical array.

Under this assumption, a copy of certain features of the *wh*-item, namely what I will call the argument licensing features (categorial features, ϕ-features, case features, etc.) can be created. The copied features remain with the *wh*-item while the original features create a new lexical item. During the derivation the *wh*-item containing the copied features and the new item containing the original features can undergo Merge separately; the *wh*-item is merged in the FR where it serves as the relative pronoun while the new item is merged in the matrix clause as the (covert) head of the FR. Thus, the *wh*-phrase is indeed part of two clauses at the same time which explains its

---

[1]Note that there are two types of FRs in German. One type uses a *wh*-phrase as a relative pronoun as exemplified in (1), while the other type uses a normal *d*-relative pronoun. Fuß and Grewendorf (2012), however, argue that this *d*-type of FRs behaves differently and should therefore not be classified as a genuine FR. In what follows, I will confine myself to the discussion of FRs with *wh*-relative pronouns.

puzzling behavior. Finally, in the post-syntactic component the copied features are deleted under a checking relation with the original features.

This account does not only provide a good basis for the analysis of the properties of FRs discussed below, but also presents a development of the covert head approach of Groos and van Riemsdijk (1981), among others, which assumes that the *wh*-phrase is part of the embedded FR and that the FR is headed by a covert element. The main problem with this approach is, however, that the special link between the covert head and the *wh*-phrase has to be explained (cf. Hirschbühler 1978). Within the present account the tight connection between these two items comes for free since the covert head is actually a part of the *wh*-item.

The paper is structured as follows: In section 2, I show that the status of the *wh*-phrase is indeed ambivalent with respect to its position. Afterwards, in section 3 I develop an account of FRs that forms the basis for the concrete analyses of the properties in section 4. Finally, in section 5 I discuss the consequences of the account with respect to the interaction of operations and compare it to previous approaches to FRs. Section 6 concludes.

2. Data

The aim of this section is to show that the *wh*-phrase seems to be part of two clauses at the same time. This is done by outlining two kinds of data: the first kind suggests that the *wh*-phrase is part of the matrix clause while the second kind suggests that it is part of the embedded clause.

In the data below I have not indicated the exact position of the *wh*-phrase. Instead, the category containing the *wh*-phrase is labeled "FR". This is done for reasons of readability should not suggest whether the *wh*-phrase is the head of this constituent or not.

2.1. The *wh*-Phrase is Part of the Matrix Clause

2.1.1. *Number Agreement*

The first fact that suggests that the *wh*-phrase is part of the matrix clause involves number agreement. Bresnan and Grimshaw (1978) observed for English FRs that a plural *wh*-phrase in an FR induces plural number agreement

in the matrix clause. This observation also holds for other languages, e.g., German.

The German data in (2) show that a plural *wh*-subject induces plural number agreement in the clause where it occurs.

(2)   *Agreement: matrix clause*
       [Welche Bücher] haben/ *hat   dir gefallen?
       which   book.PL have.PL/ have.SG you liked
       'Which books did you like?'

Now, usually plural *wh*-phrases that occur in embedded clauses do not affect the number agreement in the matrix clause. This is exemplified in (3) for indirect interrogative clauses in subject position.

(3)   *Agreement: embedded clause*
       [SClause Welche Bücher ihm gefallen], ist/   *sind unklar.
             which   book.PL him like,   be.SG/ be.PL unclear
       'It is unclear which books he likes.'

However, in case of FRs, the *wh*-phrase induces plural agreement, in contrast to indirect interrogative clauses. Note that in the sentences in (4) the entire FR is the subject of the matrix clause.

(4)   *Agreement: free relative*
       ?[FR [Rel welche Bücher] ich auch immer gelesen habe], haben/
             which book.PL I   ever   read   have   have.PL/
       *hat   mir gefallen.
       have.SG me liked
       'I liked whatever books I read.'

If it is true that number agreement in a matrix clause cannot be induced by a constituent in the embedded clause, the *wh*-phrase in (4) must be part of the matrix clause.

Note that this property of FRs can only be shown with complex *wh*-phrases in German since simple *wh*-phrases always induce singular number agreement. But the use of complex *wh*-phrases in FRs in German is often considered to be marginal and it requires the use of the particle *auch immer* 'ever', the status of which is yet to be clarified. However, the number agreement prop-

erty of FRs also occurs in languages with simple plural *wh*-phrases, as shown in (5) for Spanish.

(5)   [$_{FR}$ Quienes son   del   sur   ] son   en gran parte bajos.
          who.PL be.PL of.the south   be.PL in great part short.PL
        'Most people from the South are short.'   *(Caponigro 2003: 169)*

Again, the plural *wh*-pronoun *quienes* 'who' induces plural agreement in the matrix clause.

2.1.2.   Extraction

The second fact that suggests that the *wh*-phrase is part of the matrix clause comes from extraction data. Rooryck (1994) observed for English that nothing can be extracted out of an FR unless it is part of the *wh*-phrase. This observation also holds for German (see Ott 2011). The extraction property is illustrated for topicalization in German.

   In general, topicalization in German may cross a *wh*-island (Fanselow 1987, Müller and Sternefeld 1993), see (6-b).

(6)   *Extraction: wh-complement clause*
   a.   Ich weiß [$_{OClause}$ welche Bücher *Der Spiegel* diesen Leuten
          I   know   which books *Der Spiegel* these   people
          empfiehlt]
          recommends
          'I know which books *Der Spiegel* recommends to these people.'
   b.   ?Diesen Leuten$_i$ weiß ich [$_{OClause}$ welche Bücher *Der Spiegel* t_i
          these   people know I   which books *Der Spiegel*
          empfiehlt].
          recommends
          'As for these people, I know which books *Der Spiegel* recommends
          to them.'   *(based on Ott 2011: 188f)*

However, topicalization out of an FR results in strong ungrammaticality if the category is extracted from within the FR, see (7-b). In contrast, if the category is a part of the *wh*-phrase, topicalization is possible, see (7-c).

(7)   *Extraction: free relative*

    a.   Ich lese [$_{FR}$ welche Bücher von Jostein Gaarder auch immer
        I   read   which books by Jostein Gaarder ever
        Der Spiegel diesen Leuten empfiehlt]
        Der Spiegel these   people recommends
        'I read whatever books by Jostein Gaarder *Der Spiegel* recom-
        mends to these people.'

    b.   *Diesen Leuten$_i$ lese ich [$_{FR}$ welche Bücher von Jostein Gaarder
        these   people   read I   which books by Jostein Gaarder
        auch immer *Der Spiegel* t_i empfiehlt].
        ever   *Der Spiegel*   recommends
        'As for these people, I read whatever books by Jostein Gaarder *Der*
        Spiegel recommends to them.'   *based on (Ott 2011: 188f)*

    c.   ?Von Jostein Gaarder$_i$ lese ich [$_{FR}$ [welche Bücher t_i]
        by   Jostein Gaarder read I   welche books
        auch immer *Der Spiegel* empfiehlt].
        ever   *Der Spiegel* recommends
        'As for Jostein Gaarder, I read whatever books by him *Der Spiegel*
        recommends.'

Again, these data suggest that the *wh*-phrase is actually part of the matrix clause, because then, the contrast between (7-b) and (7-c) follows immediately: in (7-b), topicalization is excluded since it involves extraction out of a relative clause island. In (7-c), however, the topicalized PP is extracted out of an object DP in the matrix clause which is possible in German as can be seen from the data in (8).

(8)   Von Jostein Gaarder$_i$ habe ich schon   [viele Bücher t_i] gelesen.
        by   Jostein Gaarder have I   already many books   read
        'I have already read many books by Jostein Gaarder.'

Originally, the observation about extraction was made for English by Rooryck (1994: 197). The data are given in (9).

(9)   a.   I will eat [$_{FR}$ whatever the chef recommends to that person]
    b.   *This is the person [$_{RClause}$ to whom$_i$ I will eat [$_{FR}$ whatever the
        chef recommends t_i]]

c.   This is the author [$_{RClause}$ of whom$_i$ I buy [$_{FR}$ [whatever books t_i]] the NYT recommends to its readers]

Here, the type of extraction is relativization. Movement of a relative pronoun out of an FR is impossible as shown in (9-b). However, movement becomes better if the relative pronoun is extracted from the *wh*-phrase, as shown in (9-c).

To sum up, we have seen that there are indeed facts that indicate that the *wh*-phrase is located outside the embedded clause. The next subsection will deal with data that suggest the opposite, namely that the *wh*-phrase belongs to the embedded clause.

2.2.   The *wh*-Phrase is Part of the Embedded Clause

2.2.1.   *Extraposition*

The first data set shows the behavior of FRs with respect to extraposition. The observation made by Groos and van Riemsdijk (1981) is that if the FR is extraposed, the *wh*-phrase is extraposed as well, see (10). This is unexpected under the assumption that the *wh*-phrase is outside the embedded clause (as in the analysis of Bresnan and Grimshaw 1978).

In German, there is a ban on extraposing DPs. (Finite) CPs, on the other hand, are preferably extraposed. The data in (10-b) and (10-c) suggest that the *wh*-phrase *was* 'what' is part of a CP and is not the external D head of the FR. Otherwise, (10-c) should be grammatical and (10-b) ungrammatical.

(10)   *Extraposition: free relative*
a.   Ich denke, dass ich [$_{FR}$ was ich mag] essen kann.
     I   think that I   what I   like eat   can
     'I think that I can eat what I like.'
b.   Ich denke, dass ich t_{FR} essen kann, [$_{FR}$ was ich mag].
     I   think that I   eat   can   what I   like
     'I think that I can eat what I like.'
c.   *Ich denke, dass ich [was t_{FR}] essen kann, [$_{FR}$ ich mag].
     I   think that I   what   eat   can   I   like
     'I think that I can eat what I like.'

In (11), it can be seen that indirect interrogative clauses in subject position, where the *wh*-phrase clearly belongs to the embedded CP, show the same behavior as FRs when it comes to extraposition.[2]

(11)   *Extraposition: indirect interrogative clause*
  a.   [SClause Was ich mag] ist unklar.
           what I   like   is unclear
       'It is unclear what I like.'
  b.   Es ist unklar [SClause was ich mag].
       it is unclear   what I   like
       'It is unclear what I like.'
  c.   *[SClause Was t_i] ist unklar [ich mag]$_i$.
           what   is unclear I   like
       'It is unclear what I like.'

On the other hand, as shown in (12), if the *wh*-phrase is the external head of a relative clause CP, it cannot be extraposed along with the relative clause.[3]

(12)   *Extraposition:* what-*headed relative clause*
  a.   ?Wer hat [DP was [RClause das ich gern   gegessen hätte]]
       who has   what   that I   gladly eaten   would have
       weggeworfen?
       thrown away
  b.   *Wer hat weggeworfen [DP was [RClause das ich gern   gegessen
       who has thrown away   what   that I   gladly eaten
       hätte]]?
       would have
  c.   ?Wer hat [DP was t_i] weggeworfen [RClause das ich gern
       who has   what   thrown away   that I   gladly
       gegessen hätte]$_i$?
       eaten   would have

---

[2]Note that the clauses in (11) can also be understood as FRs. Under this reading, the matrix predicate *be unclear* would predicate over the things I like. The intended reading in (11) is, however, the reading of the indirect interrogative clause, i.e., the question of what I like cannot be answered clearly.

[3]Note that relative clauses headed by *wh*-phrases are generally considered to be a bit quirky in German. Nevertheless, (12-a) and (12-c) are grammatical and stand in sharp contrast to (12-b).

Thus, this argument strongly suggests that the *wh*-phrase in FR is actually part of the embedded clause.

2.2.2. *Case Matching*

The final facts that argue for a position of the *wh*-phrase in the embedded clause concern a restricted set of possible case mismatches between the embedded and the matrix clause. In general, FRs exhibit a case matching property (Bresnan and Grimshaw 1978, Groos and van Riemsdijk 1981): the *wh*-phrase must bear a case marker that fits the case assigning properties of both the matrix clause and the FR, see (13).

(13) a. Ich folge   [$_{FR}$ wem   ich vertraue]
          I   follow→DAT   who.DAT I   trust→DAT
          'I follow who I trust.'   *(Vogel 2001: 902)*
      b. *Ich folge   [$_{FR}$ wem/wen   ich bewundere]
          I   follow→DAT   who.DAT/who.ACC I   adore→ACC
          'I follow who I adore.'   *(Vogel 2001: 902)*

Only in (13-a) do both the matrix clause and the embedded clause assign dative case to the *wh*-phrase. Hence, (13-a) is grammatical. In (13-b), on the other hand, the *wh*-phrase receives accusative case from the embedded verb *bewundern* 'adore', but it receives dative case from the matrix verb *folgen* 'follow'. This case mismatch cannot be resolved and the sentence is ungrammatical, no matter which case marker the *wh*-phrase actually bears.

However, based on the case hierarchy in (14) (cf. Pittner 1991, 1995, Vogel 2001, Grosu 2003), certain case mismatches are allowed: if the case assigned by the matrix clause is higher on the hierarchy than the case assigned within the FR, the *wh*-phrase may bear the case of the FR, violating the matching condition. Importantly, if a case mismatch is allowed, the case marker on the *wh*-phrase *must* be the case assigned within the embedded clause, see (15-a) and (15-c).

(14) *Case Hierarchy*
      NOM ≫ ACC ≫ DAT (≫ GEN)

(15)   a.   [_FR Wem/*Wer   Maria   vertraut]   wird
                who.DAT/who.NOM Maria.NOM trusts→DAT is→NOM
           eingeladen
           invited
           'Who Maria trusts gets invited.'   *(Vogel 2001: 903)*
       b.   *Er zerstört   [_FR wer   ihm begegnet]
            he destroys→ACC   who.NOM him meets→NOM
            'He destroys who meets him.'   *(Vogel 2001: 904)*
       c.   [_FR Wen/*Wer   Maria.NOM mag]   wird
                who.ACC/who.NOM Maria   likes→ACC is→NOM
           eingeladen
           invited
           'Who Maria trusts gets invited.'   *(Vogel 2001: 903)*

That the *wh*-phrase must bear the case marker of the embedded clause in case of a mismatch strongly suggests that it is actually part of the embedded clause and not the matrix clause. If the *wh*-phrase was part of the matrix clause, it would remain unclear why it should appear with the morphological case marker of the embedded clause.

Before continuing, a few remarks are in order. First, there is a second exception to the matching condition that concerns morphologically syncretic forms. Thus a sentence like (16-a) is acceptable whereas (16-b) is not.

(16)   a.   Er tut   immer [_FR was   mich ärgert].
            he does→ACC always   what.NOM/ACC me   annoys→NOM
            'He always does something annoying to me.'
       b.   *Er liebt   [_FR wer   mich ärgert].
            he loves→ACC   who.NOM me   annoys→NOM.
            'He loves who annoys me.'

(16-a) is grammatical even though the matching condition is violated and the mismatch does not obey the case hierarchy. (16-b), on the other hand, is ungrammatical under the same condition. The reason for the grammaticality of (16-a) is that the morphological form *was* 'what' can mark both nominative and accusative case. In contrast, *wer* 'who' in (16-b) unambiguously encodes nominative case.

Second, for a certain group of speakers, (15-b) is actually acceptable (Pittner 1991, 1995). These speakers seem to have a slightly different case hierarchy, shown in (17).

(17)   NOM, ACC ≫ GEN, DAT.

Assuming the hierarchy in (17), the case mismatch in (15-b) should be allowed.[4]

2.3.   Interim Summary

So far, we have seen that there are arguments coming from number agreement and extraction that speak in favor of a position of the *wh*-phrase in the matrix clause, where it acts as the external head of the FR. On the other hand, there are also arguments coming from extraposition and case matching that suggest an analysis where the *wh*-phrase is part of the embedded clause.

3.   An Analysis of Free Relatives

In this section, I will develop an analysis of free relatives within a minimalist framework that accounts for the data introduced in section 2. The account makes use of the idea that there are three stages in a derivation: the lexical array (a pre-syntactic component), the syntax, and a post-syntactic component. The stages are discussed in more detail in section 5.1.

Concretely, I will assume that the copy operation that is part of Internal Merge may also apply pre-syntactically, thereby creating an additional item in the lexical array. During the syntactic derivation the structure is built. Afterwards, the copied features created pre-syntactically will be deleted in a post-syntactic component.

---

[4]It should be noted that the possibility for case mismatches in FRs has been called into question before. For example, van Riemsdijk (2006: 355) argues that syncretisms of the type shown in (16-a) are the only exception to the matching condition. He claims that other cases of mismatching, such as the grammatical versions of (15-a) and (15-c), can be traced back to the fact that German is in a state where it loses its morphological case system. Thus, speakers do not actually conceive these cases as mismatches but treat them as another instance of (16-a). Still, this explanation does not capture the hierarchy effects found in data such as (15), as it would predict (15-c) to be on a par with (15-b), which does not seem to be the case for many speakers of German.

I will start this section by outlining the main assumptions. Afterwards, I will go through a derivation of an FR and add a few remarks on the semantics of FRs within the present approach.

3.1. Assumptions

The first assumption concerns lexical items (LIs). Following standard assumptions, they are sets of features consisting of syntactic (formal), phonological and semantic features. However, I assume that these features are not part of *one* set but belong to different sets. Furthermore, I assume that there are two kinds of syntactic features that belong to different sets: features involved in argument licensing (categorial features, ϕ-features, case features etc.) and operator features (e.g. *wh*-features). An abstract schema is given in (18).

(18) $LI = \{\{arg_1, arg_2, arg_3, \dots\}, \{op_1, op_2, op_3, \dots\}, \{phon_1, phon_2, phon_3, \dots\}, \{sem_1, sem_2, sem_3, \dots\}\}$

The LIs that are relevant for a derivation are gathered in a lexical array (LA). This LA must be empty by the end of the syntactic derivation which is ensured by the application of three operations: Copy, Merge and Agree (Chomsky 1995 et seq.).[5]

Agree is a checking operation that affects features directly: a probe feature that needs a value looks for a matching goal feature that has a value and the two enter into an agreement relation whereby the probe feature is valued by the goal feature (see Chomsky 2000, 2001).

Merge is a set-building operation that acts upon sets of features: two sets α and β become the elements of a new set. The operation is formalized in (19).

(19) $Merge (\alpha,\beta) = \{\alpha,\{\alpha,\beta\}\}$

---

[5] Note that there is an ongoing trend to dismiss the copy operation and only have a simple Merge operation, see Chomsky (2013). The consequence of this assumption would be that there is a complex Internal Merge operation which comprises the copy operation. This would, however, result in a loss of the identity of External Merge and Internal Merge because different Merge operations – one with copying, the other without copying – must be defined. The simpler assumption is therefore to separate Merge and Copy. (Note that assuming Remerge (Epstein et al. 1998) instead of Copy+Merge does not help either: first of all, Remerge provides no possibility to unify Internal and External Merge since it is a separate operation by definition. Furthermore, Remerge is incompatible with the set-theoretic syntactic structure assumed in the present approach.)

For the sake of simplicity, the set structures created in (19) are represented as trees or labeled bracketing.

(20) $\{\alpha,\{\alpha,\beta\}\} = [_\alpha\ \alpha\ \beta\ ] =$

$$\alpha$$
$$\overset{\frown}{\alpha\quad\beta}$$

Following Chomsky (2001 et seq.), there are two possibilities for Merge: *External Merge* and *Internal Merge*. External Merge combines two undominated sets, that is, sets that have no supersets, i.e., lexical items or undominated complex structures. Internal Merge, on the other hand, involves Merge of a subset α of a complex structure β with β. Following Chomsky (1995 et seq.), Internal Merge involves a copy operation.

Copy, cf. (21), is an operation that precedes Internal Merge. (For the independency of Copy and Merge, see also Nunes 1995, 2004 and the discussion in footnote 5.) It creates a copy of a syntactic object. Departing from the standard assumption, I assume that the copy actually replaces the original item and that it is the original item that is merged in a new position.

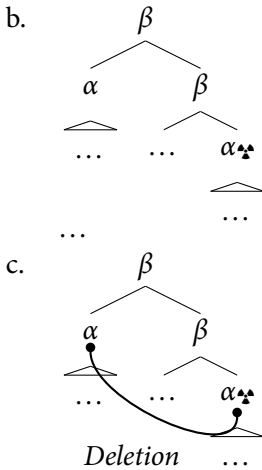(21)   Copy $(\alpha) = <\alpha,\alpha_{\clubsuit}>$

Copies must be deleted by the end of the derivation. Deletion of features applies post-syntactically for reasons of interface (especially PF) interpretability (Chomsky 1995, Nunes 2004).

In order to distinguish copied and original items, the copied item is marked by a diacritic $\clubsuit$, which is supposed to suggest that this object is unstable and has to be deleted. Deletion of a copy $\alpha_{\clubsuit}$ requires that it is c-commanded by the original item or another copied item. Conrete assumptions about the locality of post-syntactic deletion are presented at the end of section 3.2.

An abstract example of an Internal Merge operation that involves Copy is shown in (22).

(22)   Let $\beta = [_\beta\ldots\alpha]$.

    a.   (i)   Merge($\_\!\_$,β)

           (ii)   = Merge($\_\!\_$, $[_\beta\ldots\mathrm{Copy}(\alpha)]$)

           (iii) = Merge($\_\!\_$, $[_\beta\ldots<\alpha,\ \alpha_{\clubsuit}>]$)

           (iv) = Merge(α, $[_\beta\ldots\alpha_{\clubsuit}]$)

           (v)   = $[_\beta\ \alpha\ [_\beta\ldots\alpha_{\clubsuit}]]$

b.

c.

Deletion   ...

(22-a-i) shows the starting point for Internal Merge. The first argument of the operation Merge that involves a structure β still needs to be created out of β. Thus, Copy must be invoked and applies to the constituent α in (22-a-ii). In (22-a-iii), Copy has created an ordered pair consisting of the original α and the copied α. Now, in (22-a-iv) and (22-a-v), the original α can be the target for Merge while the copied α remains inside β. (22-b) shows a tree diagram of the labeled bracketing in (22-a-v). Afterwards in a post-syntactic component, the unstable occurrence of α must be deleted. This is illustrated in (22-c).

One very important assumption I make here is that the possibility for Copy is given at any time, also *before* the derivation actually starts, namely pre-syntactically in the lexical array.[6] Again, Copy may affect only sets that have supersets, i.e., parts of LIs may be copied pre-syntactically.

---

[6]Whether Merge and Agree may also operate pre-syntactically will not be discussed in what follows. Note, however, that the possibility of pre-syntactic operations is an interesting advancement to standard minimalism and facilitates the analysis of different phenomena. See, e.g., Heck (2010) for an analysis of the ban on direct recursion that is based on the possibility of pre-syntactic Agree; Assmann et al. (2012a) for an analysis of possessor advancement that builds upon the idea of pre-syntactic Agree; Georgi (2012) for a pre-syntactic deletion operation that is involved in analyzing global case splits; and Assmann (2012b) as well as Agbayani and Ochi (2007) for an approach to parasitic gaps that makes use of a pre-syntactic fission/splitting operation. See also Thomas (2013) for an approach to split ergativity that is based on pre-syntactic deletion of case features.

After an LI has entered the derivation, its internal structure is no longer accessible to Copy, i.e, lexical integrity is preserved. (This assumption might be dismissed, if theories like Chomsky 1995, Agbayani 1998, Broszewski 2003 turn out to be correct, which assume that parts of lexical items can be subject to Internal Merge, and hence Copy, also in the syntax.)

In most cases, pre-syntactic Copy of parts of LIs will lead to a crash of the derivation since there is no position available in the structure where the additional items can be merged. This is due to syntactic constraints like e.g. the Θ-Criterion (Chomsky 1981), the case filter (Rouveret and Vergnaud 1980, Chomsky 1980) or the assumption that Merge is feature-driven (e.g. Müller 2011). Hence, the additional items will remain in the LA, which causes a violation of the constraint that the LA has to be empty. However, if the LA lacks an item to begin with, Copy may create the missing item out of an existing one.

Assuming a structure of LIs as in (18), syntactic features do not form a constituent either with phonological or semantic features. Hence Copy may only affect the syntactic features of a lexical item. (If only phonological or only semantic features are copied, an element is created that is not viable in the syntax, since it has no syntactic features. Note however that Copy of phonological and/or semantic features might be of use for deriving certain PF and LF phenomena.[7])
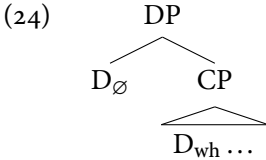
3.2.   Syntax of Free Relatives

With these assumptions in place, we can now start to derive FRs. Another example of an FR in German is given in (23).

(23)   dass alle   [FR was ich tue] gut   finden
        that everyone   what I   do   good find
        'that everyone likes what I do'

Basically, the analysis that follows is a development of the analysis first proposed by Groos and van Riemsdijk (1981). (Other versions of this analysis can be found in Grosu 1996, 2003, Citko 2004.) In this type of analysis, FRs are as-

---

[7]Furthermore, there have been proposals that Copy of syntactic features may apply post-syntactically. In order to analyze complementizer agreement, Fuß (2008) suggests that the morpheme bearing the subject agreement features is copied post-syntactically. The copy is then inserted on C which ensures that both the agreeing complementizer and the agreeing verb realize the subject's ϕ-features.

sumed to have a structure as in (24), where a CP is adjoined to a D element.[8]
This guarantees that the entire category is a DP which can be merged in an
argument position.[9]

(24)

$$
\begin{array}{c}
\text{DP} \\
\diagup\quad\diagdown \\
\text{D}_\varnothing \qquad \text{CP} \\
\diagup\diagdown \\
\text{D}_{\text{wh}} \ldots
\end{array}
$$

The main question here concerns the covert D head. It is often assumed that
this head comes directly from the lexicon. Furthermore, the data presented
in section 2 suggest a special, very close, link between the covert D head and
the *wh*-phrase, a link not found in headed relative clauses. This tight relation
between the *wh*-phrase and the covert head basically leaves two possibilities
within this type of approach: first, one could assume that the covert head is
massively underspecified and needs to Agree with the *wh*-phrase in all its fea-
tures (even the categorial feature, see footnote 9). This begs the question as to
why the lexicon should provide such a massively underspecified covert head
to begin with (see also the discussion in section 5.2.2.1).

Therefore a second possibility is pursued here: the covert head is a copy of
the *wh*-phrase, more exactly, of a part of it.

The main assumption that the derivation of FRs is based on is that the lex-
ical array (LA) of a sentence containing an FR contains only one *wh*-phrase.
This *wh*-phrase can only be merged either inside the FR, where it fills an argu-
ment position and satisfies the *wh*-feature of the embedded C, or in the matrix
clause, where it fills an argument position.

---

[8]Assuming bare phrase structure (Chomsky 1995), the D element is a head and a phrase at
the same time.

[9]Note that there are also FRs of other categories, e.g., AP-FRs (cf. mainly Grosu 1996, 2003).
A German example is given in (i).

(i)   Ich zahle [FR wie viel   du bezahlst hast].
       I   pay   how much you have   paid.

In the present account, the external head of the FR results from pre-syntactic Copy that applies
to the *wh*-phrase. Essentially, the categorial feature of the *wh*-phrase is also copied. Thus, the
external head of an AP-FR is expected to be of category A, just like the *wh*-phrase.

(25)   *Working hypothesis*[10]
        The lexical array of a sentence containing an FR is deficient and does
        not provide enough LIs to guarantee a converging derivation.

Assuming that one DP can only fill one argument position (presumably for
case and theta-role reasons), one possibility to ensure that both clauses have
enough arguments is by Copy applying in the LA. In this way an additional
item is created that may be merged in another argument position. For reasons
outlined below, Copy has to affect the *wh*-item. The feature structure of the
wh-item *was* ('what') is given in (26).

(26)   *was* = {{D,ϕ:3sg,case:__},{wh}, PHON, SEM}

An important point about the structure in (26) is that certain features, namely
the argument licensing features {D, ϕ, case, …} form a constituent to the ex-
clusion of the operator features, e.g., the *wh*-feature. The intuition behind this
structure is that there are certain features that are necessary for a constituent
in order to fill an argument position, e.g., the categorial feature to satisfy the se-
lectional needs of the verbal projection, ϕ-features to satisfy the needs of the
functional heads v and T, and case features to pass the case filter (Rouveret
and Vergnaud 1980, Chomsky 1980). The *wh*-feature, on the other hand, is an
operator feature that is not needed to fill an argument position but rather to
satisfy the needs of the C head of the clause. In that sense it does not classify
as an argument licensing feature and is therefore part of a different feature set.

By definition, Copy can only affect subsets. Now, given the structure in
(26), there are various possibilities as to how Copy can apply to *was*. There is,
however, only one possibility that will lead to a converging derivation.

The sets comprising the semantic and the phonological features cannot be
copied for reasons outlined above in section 3.1: copying would create an LI
that is not viable in the syntax since it does not have any syntactic features.

The next option consists of copying the set containing the operator features.
In this case, the item created would not have features that allow it to be merged
in an argument position since it would lack the necessary argument licensing
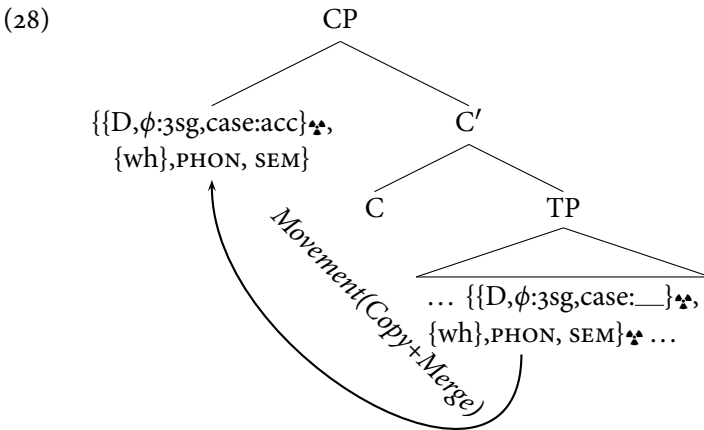features.

---

[10] See also Nunes (1995, 2004) for the idea that the structure of the lexical array is not subject
to any constraints (as originally proposed in Chomsky 1995). Rather lexical arrays are more or
less arbitrary collections of lexical items, which gives rise to the possibility of deficient lexical
arrays.

The only option that leads to a converging derivation is therefore to copy only the set containing the argument licensing features.
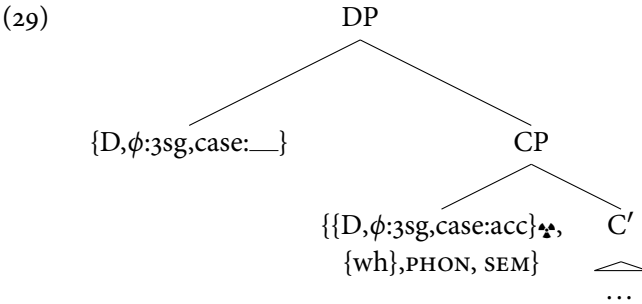
Thus the operation applying in the first stage of the derivation – the pre-syntactic component – is Copy of the argument licensing features of the *wh*-item. This is illustrated in (27).

(27)   *Pre-syntactic Copy*
       {Copy({D,ϕ:3sg,case:__}), {wh}, PHON, SEM}
       = {<{D,ϕ:3sg,case:__},{D,ϕ:3sg,case:__}♠▴>, {wh}, PHON, SEM}
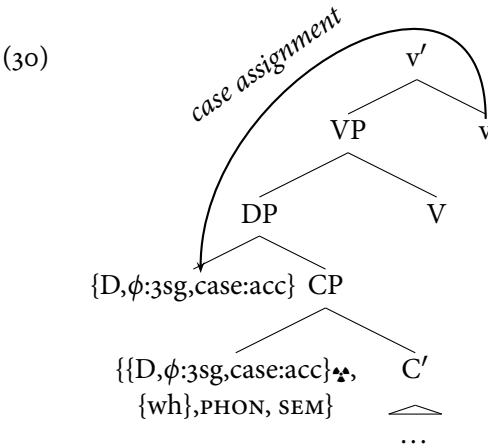   = {D,ϕ:3sg,case:__}, {{D,ϕ:3sg,case:__}♠▴, {wh}, PHON, SEM}

When the syntactic derivation starts, the embedded clause is built first. The *was*-item that contains the *wh*-feature must be merged in this clause in order for C to check its *wh*-feature. In its base position, *was* receives accusative case from the embedded v. Movement of *was* to Spec-C leaves a copy behind that must be deleted later on.

(28)

The set {D, ϕ, case} created by pre-syntactic Copy can now be merged with this CP, obtaining the structure in (24), see (29). Note that the D head does not contain any phonological features and is thus covert.

(29)

```
                              DP
                    ┌─────────┴─────────┐
          {D,φ:3sg,case:___}            CP
                                 ┌───────┴───────┐
                        {{D,φ:3sg,case:acc}♣,    C′
                         {wh},PHON, SEM}        ___
                                                ...
```

Afterwards, the entire DP can be merged as an argument of the matrix clause and the covert D head receives case by the matrix v, see (30).

(30)


```
                    case assignment        v′
                              ┌──────────────┴──────┐
                             VP                      v
                      ┌───────┴───────┐
                     DP                V
              ┌───────┴───────┐
     {D,φ:3sg,case:acc}      CP
                       ┌───────┴───────┐
              {{D,φ:3sg,case:acc}♣,     C′
               {wh},PHON, SEM}         ___
                                       ...
```

Finally, after the syntactic derivation is terminated and the covert head of the FR has received case, post-syntactic deletion of the copied feature set on the *wh*-phrase applies under c-command and identity with the set containing the original features.[11]

---

[11]In the context of FRs, it becomes clear why deletion of copies cannot apply in the syntax already: assuming that the copied feature set on the *wh*-phrase can only be deleted if all the features on the original feature set are valued, case assignment must precede deletion. If deletion was a syntactic operation, its application in the configuration in (30) would be counter-cyclic because deletion would affect only the DP-cycle although the DP is already dominated by another cycle. Assuming that Strict Cyclicity (Chomsky 1973) holds for all syntactic operations, deletion of copies must then apply post-syntactically.

(31)
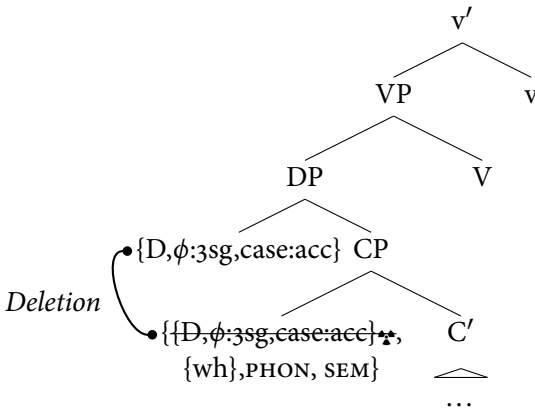
Up to now, we have achieved the following result: pre-syntactic Copy of the argument licensing features of the *wh*-item creates a new LI in the lexical array that is needed to enable a converging syntactic derivation. Post-syntactically, the copied features are deleted again.

The account of FRs proposed here can provide answers to the three main questions raised by this construction. First, the question why a clause can occur in a position reserved for non-clausal arguments is answered by assuming the structure in (24) where the clause is the complement of a covert D head. The entire structure is a DP and can therefore occur in DP positions.

Second, the head of the FR must be covert because it is an LI created by pre-syntactic Copy. Since the phonological features cannot undergo pre-syntactic Copy for theory-internal reasons, the new LI, which serves as the external head later on, is covert.
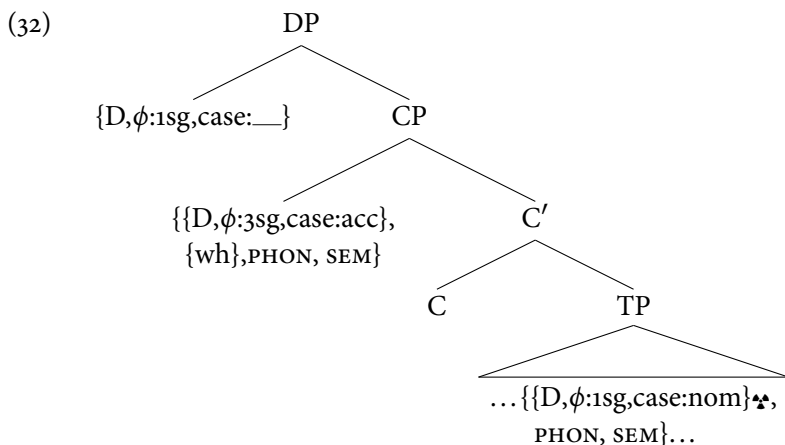
Third, the close link between the *wh*-phrase and the covert head arises from the fact, that the covert head is made up of features from the *wh*-head. Thus, the features and their values have to coincide.

Before I proceed with the analysis of the properties of FRs, one question remains to be answered: why must pre-syntactic Copy affect the *wh*-phrase? The reasons are the following: First, since it is a D item that is actually missing, Copy of any other category but D would not bring the required result. Thus, we have already narrowed down the possibilities for pre-syntactic Copy to the D items in the lexical array.

Furthermore, Copy of some D item that appears in the matrix clause would also lead to a crashing derivation because in this case, the copied features would be higher in the structure than the original features, i.e., exactly the

opposite of the structures above. Under the assumption that the original features have to c-command the copied features in order for deletion to apply, the derivation would crash due to undeleted copied features.

Now, there are not many D elements left. In a transitive FR-clause like (23) it boils down to two D items: *was* 'what' and the pronoun *ich* 'I'.[12] Now, what would exclude a structure as in (32) where Copy applied to *ich*?

(32)

DP
├── {D,ϕ:1sg,case:__}
└── CP
    ├── {{D,ϕ:3sg,case:acc}, {wh},PHON, SEM}
    └── C′
        ├── C
        └── TP
            └── …{{D,ϕ:1sg,case:nom}♣, PHON, SEM}…

A way to exclude (32) would be to assume that spell-out applies cyclically and that copies are tolerated only for a short time in the post-syntactic component. Concretely, (32) can be ruled out as follows. Assume, following Chomsky (2000, 2001), Heck and Zimmermann (2004), Svenonius (2004), that phases are vP, CP and DP. As soon as a phase is built, its complement is sent to Transfer. In order to escape spell-out, categories must move to the edge of the current phase. These assumptions ensure that at most one phase boundary may intervene between a copy and its antecedent (another copy or the original category). In other words, the antecedent of a copy enters the post-syntactic component always on the next spell-out cycle. Assuming further, that this is a necessary condition for deletion to apply, the structure in (32) can be excluded. In (32), the subject inside the TP contains a copied feature set. Deletion of this copy may only apply if the CP contains a potential antecedent for it. This, however, is not the case in (32). Put differently, the problem with (32) is that

---

[12] For the sake of concreteness I assume that pronouns are simple D heads.

it looks as if movement would have skipped the CP phase (even though no actual syntactic movement took place).[13]

Having excluded all other options, the only converging derivation results from Copy of the *wh*-item that moves to Spec-C of the embedded clause for independent reasons. In sum, the options for pre-syntactic Copy are highly limited by the conditions for converging derivations.

In the next part of this section, I will show that the syntactic structure created so far is compatible with a standard semantics for free relatives.

3.3.   Semantics of Free Relatives

The syntactic structure of FRs derived above is compatible with standard semantic analyses of FRs. Following Caponigro (2003), Jacobson (1995) (cf. also Grosu and Landman 1998, Grosu 2003), I assume that FRs are semantically like DPs in that both denote the maximal entity (Link 1983) described by a predicate.

Furthermore, I assume that the semantic type of a lexical item depends on its features: *wh*-phrases, e.g., have a *wh*-feature and are therefore of type $<<e,t>,<e,t>>$ (Caponigro 2003).[14]

The λ-representation of the *wh*-item *was* is given in (33) (cf. Caponigro 2003).

(33) $\{\{D,\phi:3sg,case:\_\_\}_{\clubsuit},\{wh\},\text{PHON}, \text{SEM}\}$
        $<<e,t>,<e,t>>: \lambda P \lambda x[-\text{anim}'(x) \land P(x)]$

Following Caponigro (2003), questions are predicate abstracts. That is, the *wh*-item *was* "applies to a set of entities to give back all and only the entities of that set that are inanimate" (Caponigro 2003: 58).

---

[13]Of course, this situation can only arise if Copy has applied pre-syntactically. Then, the copy and the original object can enter the derivation at any point and no locality conditions can be imposed on them in the syntax, in contrast to applications of Copy that precede syntactic movement.

[14]At this point, the question arises as to what extent the semantics of a lexical item depends on its semantic features and what semantic features actually are. Following Chomsky (1995: 230), semantic features are to be understood as descriptive features that constitute the result of a semantic decomposition of the meaning of the lexical item. (Chomsky gives the example *airplane* that has a semantic feature [artifact].) Thus, only non-functional elements have true semantic features. The semantic type and the formal parts of the λ-representation of an LI, on the other hand, depend on the formal, i.e, the syntactic features.

The semantic operator σ that returns the maximal entity of this set is of type $<<e,t>,e>$, i.e., basically the same type as a simple D head. Such a simple D head is indeed available in the theory above due to Copy: it is the head of the DP dominating the embedded CP. This head has the feature specification of a simple D head, a semantics for which is given in (34).

(34) $\{D,\phi:3sg,case:\_\_\}$
       $<<e,t>,e>:\lambda P \sigma x[P(x)]$

In (34), the σ-operator takes a set and returns its maximal entity. With these assumptions in mind, the meaning of FRs can be computed compositionally on the basis of the structure in (35).

(35)
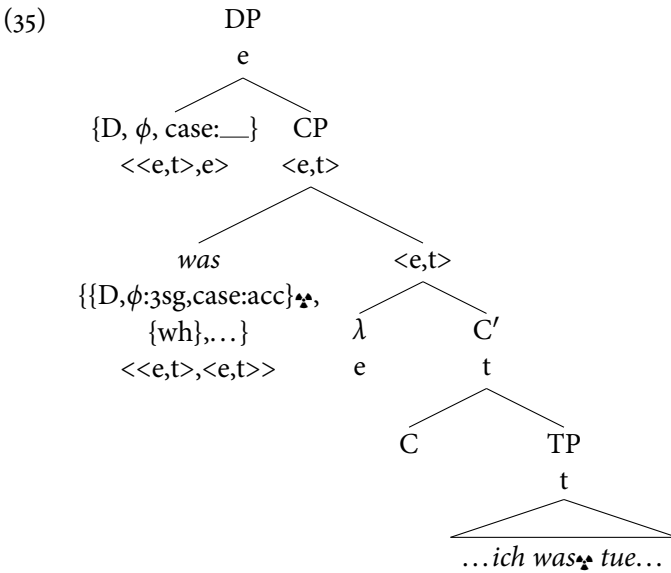
The crucial point is what happens in the CP and the DP. Movement of *was* to Spec-C has created a λ-abstract $\lambda x_1[do'(x_1)(\text{SPEAKER})]$. Now, the *wh*-item *was* applies to the set described by this λ-abstract and we end up with a form $\lambda x_1[-anim'(x_1) \wedge do'(x_1)(\text{SPEAKER})]$. Finally the D head applies to this set and returns its maximal entity: $\sigma x_1[\lambda x_1[-anim'(x_1) \wedge do'(x_1)(\text{SPEAKER})]]$.

As can be seen, the analysis of Caponigro (2003) is fully compatible with the syntactic structure of FRs as proposed here. Furthermore, the emergence

of the σ-operator – a point not explained in Caponigro (2003) – is captured: it is the result of pre-syntactic Copy.[15]

4. Deriving the Behavior of the *wh*-Phrase

In this section, I will show how the properties of FRs introduced in section 2 can be analyzed within the account outlined in section 3.

4.1. The *wh*-Phrase is Part of the Matrix Clause

4.1.1. Number Agreement

The first property described in section 2.1.1 concerned number agreement. The observation was that a plural *wh*-phrase in an FR induces plural number agreement in the matrix clause. Within the present account the data can be derived as follows. Let's first have a look at a simple *wh*-phrase, as illustrated by the Spanish data in (36), repeated from (5).

(36) [FR Quienes <u>son</u> del   sur   ] <u>son</u> en gran parte <u>bajos</u>.
          who.PL be.PL of.the south   be.PL in great part   short.PL
       'Most people from the South are short.'

A plural number feature is part of the ϕ-features of *quienes* 'who'. Hence, after pre-syntactic Copy there are two plural number features, one on the item *quienes* and one on the newly created D head. The pre-syntactic Copy operation is shown in (37).

(37) a. *quienes* = {{D,ϕ:3pl,...}, { wh}, PHON, SEM}
       b.   {Copy({D,ϕ:3pl,...}), {wh}, PHON, SEM}
            = {D,ϕ:3pl,...}, {{D,ϕ:3pl,...}♠, {wh}, PHON, SEM}

Thus, plural number agreement is supposed to be possible in the FR as well as in the matrix clause. The Agree relations are illustrated in (38).

---

[15]Note that the analysis of Caponigro (2003) was assumed here for the sake of concreteness. Since the resulting syntactic structure of FRs is quite a standard, very widespread one, any other semantic analysis that is based upon the structure in (24) should be compatible with the syntactic analysis of the present approach.

(38) [$_{vP}$ [$_{DP}$ {ϕ:3pl,...} [$_{CP}$ {{ϕ:3pl,...}♠,{wh}, ...} ...son...]] ...son...]

number agreement
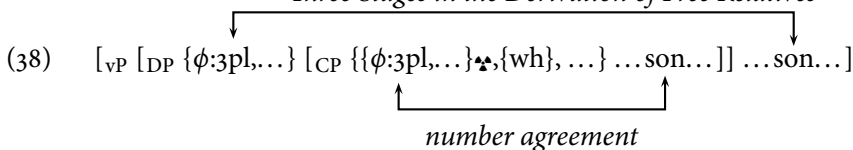
Things are more complicated with the complex *wh*-phrase in German in (39), repeated from (4).

(39) ?[$_{FR}$ [$_{Rel}$ welche <u>Bücher</u>] ich auch immer gelesen habe], <u>haben</u> mir
         which book.PL I   ever   read   have have.PL me
      gefallen.
      liked
      'I liked whatever books I read.'

In general, there are two possibilities to extend the analysis in order to derive the German data. First, there are two forms of *welch* ('which'): a singular form selecting for a singular NP and a plural form selecting for a plural NP. Under these assumptions, the same analysis as sketched in (38) can account for (39).

   The second, perhaps more attractive, possibility is that *welch* agrees with the NP in its number feature. Then, both the original and the copied number feature must agree with the number feature of the NP which is merged in the FR. The analysis is sketched in (40).

(40) [$_{DP}$ {D,ϕ:3pl,...} [$_{CP}$ [$_{DP}$ {{D,ϕ:3pl,...}♠,...} [$_{NP}$ Bücher$_{pl}$]]]]

number agreement
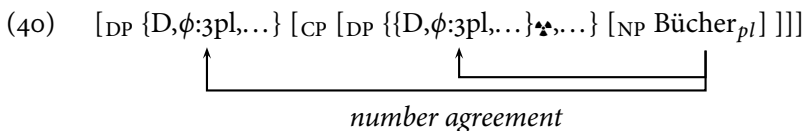
4.1.2. Extraction

The second set of data involved extraction. It was shown in section 2.1.2 that topicalization out of an FR is impossible if the category is a part of the FR. On the other hand, if the category is a part of the *wh*-phrase, topicalization is possible. The crucial data are repeated in (41).

(41) a. *Diesen Leuten$_i$ lese ich [$_{FR}$ welche Bücher von Jostein Gaarder
          these   people read I   which books by Jostein Gaarder
          auch immer *Der Spiegel t$_i$* empfiehlt].
          ever   *Der Spiegel*   recommends
          'As for these people, I read whatever books by Jostein Gaarder *Der Spiegel* recommends to them.'

b.   ?Von Jostein Gaarder$_i$ lese ich [$_{FR}$ [welche Bücher t_i]
        by   Jostein Gaarder read I   welche books
        auch immer *Der Spiegel* diesen Leuten empfiehlt].
        ever   *Der Spiegel* these   people recommends
        'As for Jostein Gaarder, I read whatever books by him *Der Spiegel*
        recommends to these people.'

Now, whatever prohibits extraction out of a relative clause, rules out extraction from FRs since in this theory, they are relative clauses. Hence, (41-a) is expected to be ungrammatical.

In those cases where extraction out of an FR seems to be possible, I claim that extraction actually proceeds from outside the FR. The PP complement can choose to be merged either to the covert head outside the FR or inside the overt *wh*-phrase. If it is merged outside the FR, it can be topicalized.[16]

(42) … [$_{DP}$ [$_{DP}$ {D,ϕ:3pl,…} [$_{PP}$ von J.G.]] [$_{CP}$ {{D ϕ:3pl♣,…},…},…}.. α…]]

            Topicalization
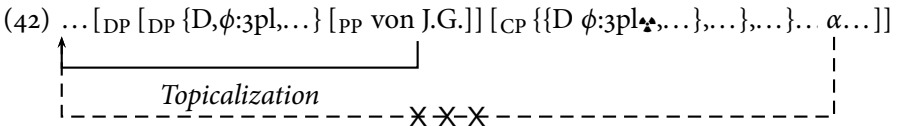
A technical way to capture this idea would be to assume that the feature selecting the PP is not an intrinsic feature but an optional feature that is added to the item in the lexical array.[17] Then, a feature [uP] that is responsible for taking a PP complement can end up either on an item to be merged in the *wh*-phrase inside the FR or it can end up on the newly created D head.[18]

---

[16]Note that if the PP is merged with the covert head, the relative clause CP cannot be directly merged with the D head. Hence, the D head does not c-command the *wh*-phrase, which is a necessary condition for post-syntactic deletion to apply. It thus seems necessary to relax the conditions for deletion and assume that m-command is sufficient. As far as I can see, this assumption does not lead to any false predictions.

[17]This is a slight divergence from Chomsky (1995) who claims that optional features are "added as the LI enters the numeration"[p.231].

[18]That D heads may take PP complements is independently verified by sentences such as (i) where the overt D head *das* is combined with a *von*-PP.

(i)   Ich habe [$_{DP}$ das [$_{PP}$ von ihm]] gelesen.
       I   have   that   by him   read.
       'I have read that one by him.'

4.2. The *wh*-Phrase is Part of the Embedded Clause

4.2.1. *Extraposition*

The extraposition data in section 2.2.1 showed that if the FR is extraposed, the *wh*-phrase is extraposed as well (Groos and van Riemsdijk 1981), see (43).

(43) a. Ich denke, dass ich t_{FR} essen kann, [$_{FR}$ was ich mag].
         I   think that I   eat   can   what I   like
         'I think that I can eat what I like.'
      b. *Ich denke, dass ich [was t_{FR}] essen kann, [$_{FR}$ ich mag].
         I   think that I   what   eat   can   I   like
         'I think that I can eat what I like.'

In the present account, the grammaticality of (43-a) and the ungrammaticality of (43-b) are derived because of the restrictions for pre-syntactic Copy of phonological features. Remember that only syntactic features may undergo pre-syntactic Copy. The result of this restriction is that the newly created D head does not possess any phonological features. The phonological features are only available on the *wh*-item *was*, that is, the phonological features are part of the embedded CP. This in turn means that the *wh*-phrase must be pronounced within the CP, i.e., if the CP is extraposed, the *wh*-phrase must be extraposed as well.[19]

4.2.2. *Case Matching*

Finally, it was shown in section 2.2.2 that certain case mismatches with FRs are allowed: if the case assigned by the matrix clause is higher on the case hierarchy than the case of the FR, the *wh*-phrase may bear the case of the FR, violating the matching condition. The case hierarchy is repeated in (44).

---

[19] The fact that FRs can be extraposed seems to be incompatible with the condition for post-syntactic feature deletion as outlined in section 3.1 and 3.2. In the extraposed position the unstable copied feature of the *wh*-phrase are not c-commanded by the covert D head. Then, deletion should not be possible, which leads to a crash of the derivation. It thus seems necessary that feature deletion applies before extraposition. This could be indeed the case if extraposition is a post-syntactic operation itself (cf. Chomsky 1995, Truckenbrodt 1995, Göbbel 2007, Wurmbrand 2007, Chomsky 2008; see Inaba 2005 for German), which would also explain why it does not behave like typical syntactic movement (see e.g. Fox and Nissenbaum 1999).

(44) *Case Hierarchy*
       NOM $\gg$ ACC $\gg$ DAT ($\gg$ GEN)

These data can be analyzed within the present account of FRs under the following assumptions: First, case features are rather case slots to which case features are added (cf. also Assmann et al. 2013, where the same assumption for analyzing the case split on possessors in Udmurt is made).

Second, cases are decomposed in a way that a case higher on the case hierarchy is a superset of a case lower on the hierarchy (cf. Béjar and Řezáč 2009; for similar ideas, see Trommer 2006, 2008). The case decomposition is given in (45).[20]

(45) *Case decomposition*
       NOM $[\alpha,\beta,\gamma] \supset$ ACC $[\alpha,\beta] \supset$ DAT $[\alpha]$

Furthermore, the identity condition of deletion will be revised slightly: in order for a copied feature set to delete, it must be a subset of the original feature set, that is, all features in the copied feature set must have a matching feature in the original feature set.

Now, if the case assigned in the FR is higher on the hierarchy, i.e., if it consists of more features than the case assigned in the matrix clause, the copied feature set contains more case features than the original feature set. In this case deletion cannot apply, as shown in (46).

(46) *Er zerstört [$_{DP}$ D   [$_{CP}$ wer   ihm begegnet]]
       he destroys $\emptyset_{\{\alpha,\beta\}}$ who$_{\{\alpha,\beta,\gamma\}}$♣ him meets
       └ α,β ┘↑ └××┘ └ α,β,γ ┘

In (46), the *wh*-item *wer* receives nominative from the functional head T in the embedded clause. Thus the copied case slot contains the features $[\alpha,\beta,\gamma]$♣. The original case slot on the external D head, on the other hand, receives accusative, i.e. $[\alpha,\beta]$. Then, however, there are features in the copied feature set on the *wh*-phrase for which there are no matching features on the original feature set. Thus, deletion cannot apply, which leads to a crash of the derivation.

---

[20]Note that the decomposed case features are abstract in (45). In what follows, I will refrain from exchanging the abstract case features by concrete features since it is not important for the idea of the analysis.

In contrast, if the case assigned in the FR is lower on the hierarchy, i.e., consists of fewer features than the case assigned in the matrix clause (or if both cases are identical), all case features on the copy inside the FR have a matching feature on the original and deletion can apply, see (47) and (48).

(47)   Ich folge [$_{DP}$ D   [$_{CP}$ wem   ich vertraue]]
         I   follow $\emptyset_{\{\alpha\}}$ who$_{\{\alpha\}}$♠ I   trust

(48)   [DP D   [$_{CP}$ Wen/   *Wer   Maria mag]] wird eingeladen
          $\emptyset_{\{\alpha,\beta,\gamma\}}$ who$_{\{\alpha,\beta\}}$♠/ $_{\{\alpha,\beta,\gamma\}}$♠ Maria like   is   invited

The trivial case is illustrated in (47). Here the case features assigned within the embedded clause and the case features assigned within the matrix clause are identical. Thus the original and the copied case bear an identical feature set. Consequently, deletion of the copied feature set is possible.

In (48), the copied case slot in the embedded clause bears fewer features than the original case slot on the external D head. More concretely, in (48), the *wh*-phrase *wen* bears the case features [α,β]♠ while the covert external D head bears the case features [α,β,γ]. Since both [α] and [β] find a matching feature on the D head, deletion of the copied feature set is possible and the derivation will converge. The case hierarchy involved in the mismatches is therefore correctly derived.[21]

---

[21]The analysis developed so far has nothing to say about the syncretism exception to the hierarchy effects, illustrated in (16) in section 2.2.2. A possible extension of the account that could derive the facts would consist in assuming that the morphological operation Impoverishment can apply in the syntax already (Keine 2010, Doliana 2013). The relevant rule with respect to FRs would be that in the context of inanimate nominative *was*, the nominative feature [γ] deletes and *was* bears only the accusative features [α,β].

(i)   [γ] → $\emptyset$/__[+wh–anim]

Consequently, in a context like (16-a), repeated below in (ii), [γ] deletes (marked as ($\gamma\!\!\!/$) in (ii)) and a fatal case mismatch in FRs is circumvented, since the covert head and overt *was* possess an identical case feature set.

(ii)   Er tut   immer [$_{DP}$ D   [$_{CP}$ was   mich ärgert]].
         he does always $\emptyset_{\{\alpha,\beta\}}$ what$_{\{\alpha,\beta,(\gamma\!\!\!/)\}}$♠ me   annoys

Finally, note that the *wh*-phrase must always be pronounced with the case of the FR. Since the phonological features are a part of the *wh*-phrase in the embedded CP, they must realize the case assigned in this clause. Hence, the ungrammatical version of (48) is correctly excluded.

4.3.   Interim Conclusion

Free relatives are puzzling since their relative pronouns (*wh*-phrases) seem to be simultaneously part of two sentences. The special behaviour of the *wh*-phrase is derived by making use of timing of operations: a part of the *wh*-phrase, namely the part containing the argument licensing features, is copied pre-syntactically in the lexical array. Then, the *wh*-phrase containing only a copy of its argument licensing features is merged inside a relative clause CP, where it is moved to Spec-C. The original argument licensing features are merged as the external head of the relative clause. The copied features on the *wh*-phrase must be deleted post-syntactically under c-command and feature identity.

Under these assumptions the puzzle concerning the ambivalent status of the *wh*-phrase is resolved: due to pre-syntactic Copy, certain features of the *wh*-phrase are available not only in the embedded clause but also in the matrix clause, namely exactly those features that are involved in the properties that suggest that the *wh*-phrase is located outside the FR, which are, e.g., ϕ-features and subcategorization features. However, all the other features of the *wh*-phrase are only available inside the FR, e.g., phonological features. Exactly those properties which suggest that the *wh*-phrase is inside the FR can be traced back to these features.

The restricted occurrences of case mismatches were derived by decomposing the case features in a way that models the case hierarchy: a case x which is lower on the case hierarchy than a case y is represented by a subset of the case features of y. Under this assumption, the hierarchy effects follow because the case features of the copied case slot must be in a subset relation $\subseteq$ with the features of the original case slot for deletion to apply, which is only given if the case assigned by the matrix clause is higher than or equal to the case assigned in the embedded clause.

_____

5.   Discussion

5.1.   The Three Stages in More Detail

The main clue of the present account of FRs is that it distributes the operations which are necessary for the derivation to converge among three different components; thus, the derivation proceeds in three different stages. The first stage subsumes operations that apply in the lexical array, before the actual syntactic derivation starts. The second stage comprises the operations in the syntactic derivation. Finally, in the third stage, post-syntactic operations apply.

Due to this division of labor into three components which are strictly ordered, the operations that apply in these components are ordered as well. The order of the operations that are relevant for the present theory of FRs is given in (49).

(49)   *Pre-syntactic*   *syntactic*   *post-syntactic*
       Copy   Copy, Merge, Agree   Deletion (of Copies)

The order of the application of these operations that result from the distribution among different components leads to certain patterns of interaction. The main schema is the following (cf. Pullum 1979): operations that apply earlier in the derivation can affect the application of operations that apply later (feeding, bleeding, cf. Kiparsky 1971, 1976), but not vice versa (counter-bleeding, counterfeeding, ibid.).

In terms of the present approach this means that pre-syntactic Copy can bleed or feed operations in the syntax (Agree, Merge), and operations in the syntax can feed or bleed post-syntactic deletion. On the other hand, operations that apply later cannot affect operations that apply earlier. This leads to potential instances of counter-bleeding and counter-feeding. In what follows, I will summarize the most interesting interactions of operations that arise from the present theory.

Let us first have a look at the transparent interactions. The most obvious example for such an interaction concerns the effects of pre-syntactic applications of Copy. Since pre-syntactic Copy creates a new lexical item, it feeds Merge of this new lexical item as the covert head of the relative clause as well as Agree relations that involve this covert D head (e.g. number agreement in the matrix clause, cf. section 2.1.1 and 4.1.1). Furthermore – indeed a more

trivial interaction – pre-syntactic Copy, which creates unstable features, feeds post-syntactic deletion of these features.

But there are also opaque interactions of operations in the present account of FRs. A very interesting kind of counter-feeding interaction involves the case matching property of FRs. In section 2.2.2 it was shown that certain mismatches are allowed while others are not. Leaving the decomposition of case features as pursued in section 4.2.2 aside for a moment, we have an instance of counter-feeding: usually, valuation of the case feature on the covert head feeds deletion of the case features (by assumption unvalued features cannot be deleted; cf. section 3.1). However, in contexts where a case mismatch is not allowed (e.g., (13-b)) this feeding relation is not given; hence, we have a counter-feeding relation. The special nature of this counter-feeding relation is that it cannot be resolved by re-ordering the operations because syntactic case assignment must in any case precede post-syntactic deletion. Instead, counter-feeding is resolved by decomposing the case values as in section 4.2.2 where a case higher on the hierarchy in (14) consists of more features than a lower case. If the *wh*-phrase in the embedded CP receives a high case, while the covert head receives a low case, the unstable copy of the argument licensing features on the *wh*-phrase cannot be deleted since there are not enough matching case features in the original feature set on the covert head.

This survey of interaction types ends the discussion of the three stages in the derivation of free relatives. The discussion has shown that the order of operations arises from a standard architecture of grammar, where the pre-syntactic component precedes the syntactic one, which in turn precedes the post-syntactic component. Since this ordering of grammatical components is strict, the order of operations is strict as well. Thus, re-ordering is not an option for solving all cases of opaque interactions. Instead, reanalyzing the data in a way that a counter-feeding relation does not arise in the first place is in line with the present account of FRs.

5.2.   Previous Analyses

In this section I will summarize previous approaches to FRs. All of them are deficient in the sense that they cannot derive some of the properties of FRs.

5.2.1. External-Head Accounts

I start this overview with a discussion of one of the oldest formal approaches to free relatives. Due to facts like the ones presented in section 2.1, Bresnan and Grimshaw (1978) claimed that the *wh*-phrase is the external head of a relative clause which contains a pronoun that is bound by the *wh*-phrase and undergoes a process of *Controlled Pro-Deletion*. The rough structure is depicted in (50).[22]

(50)   DP
    ╱ ╲
  wh$_i$ CP
         ╱╲
     …~~pro$_i$~~ …
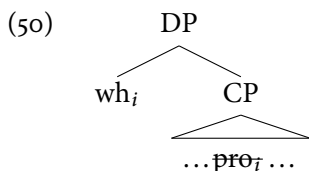
The main problem with this approach is certainly its conflict with the extraposition data from languages like German, as first pointed out by Groos and van Riemsdijk (1981) and discussed in section 2.2.1 and 4.2.1. Within these approaches, it must be possible to extrapose DPs in order to account for the data. However, DP extraposition is not attested otherwise in German. Thus, the approach cannot derive the empirical facts correctly.

5.2.2. Comp Accounts

5.2.2.1. Groos and van Riemsdijk (1981)

In order to derive the extraposition data, Groos and van Riemsdijk (1981) propose that the *wh*-phrase is inside a relative clause that is headed by an empty category.[23] The analysis is sketched in (51).

(51)   DP
    ╱ ╲
   Ø   CP
      ╱ ╲
    wh   C'
        ╱╲
      … t …

---
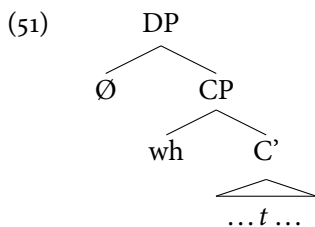
[22] See also Bresnan (1973), Daalder (1977), Larson (1987) for other head-external approaches.

[23] The earliest version of this approach was pursued by Kuroda (1968). Other variants can be found in Hirschbühler and Rivero (1981), Harbert (1983), Suñer (1984), Grosu and Landman (1998), Grosu (2003), Caponigro (2002), Gračanin-Yuksek (2008) among others.

Even though the structure in (51) is identical to the structure of FRs in the present approach, former proposals as to how this structure arises lack an explanation for the special relationship between the covert head and the *wh*-phrase. In principle, there are two ways to model this relation: (i) excessive Agree which basically copies all the necessary features from the *wh*-phrase onto the covert head (Grosu 2003 notes that the covert head must even agree in categorial features, since there are also adjectival free relatives; see also footnote 9) and (ii) the approach developed above where the covert head emerges directly from the *wh*-phrase and has the necessary features to begin with.

The potential problem of the first option is the following: it remains unclear why the covert head in free relatives behaves different from the overt head in light-headed relatives, especially when it comes to case matching. In contrast to free relatives, light-headed relatives freely allow case mismatches. This can be seen in (52).

(52)   a.   *Ich folge   Ø [_FR wem/wen   ich bewundere]
            I   follow→DAT   who.DAT/who.ACC I   adore→ACC
            'I follow who I adore.'
       b.   Ich folge   dem   [_Rel den   ich
            I   follow→DAT that.one   who.DAT/who.ACC I
            bewundere]
            adore→ACC
   'I follow the one who I adore.'

In the empty-head Comp account, the difference between the two structures in (52) is that the external head of the relatives clause is covert in (52-a), but overt in (52-b). However, as (52) suggests, this cannot be the only difference. There must be a special construction-specific constraint on FRs that force the covert head to check also its case features with the *wh*-phrase. It is far from clear how such a specific constraint could be justified.

If, on the other hand, the covert head in (52-a) does not come from the lexicon like the overt head in (52-b), but results from copying (part of) the *wh*-phrase, the close link between the covert head and the *wh*-phrase in (52-a) arises naturally. At the same time, the difference between free relatives and light-headed becomes clear immediately: in light-headed relatives, the external head does not emerge from the relative pronoun but is a separate lexical

item. Thus, the relation between the overt head and the relative pronoun is supposed to be different.
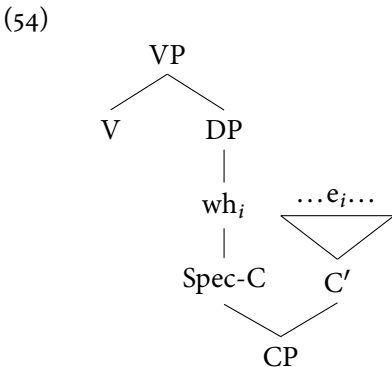
5.2.2.2. Rooryck (1994)

The next approach, I will discuss was first proposed by Rooryck (1994).[24] Similarly to Groos and van Riemsdijk's (1981) account, the *wh*-phrase is located inside the embedded CP. In contrast to that approach, however, the CP is directly merged in an argument position in the matrix clause. The structure is shown in (53).

(53)

$$
\begin{array}{c}
\text{CP} \\
\diagup\diagdown \\
\text{wh} \quad \text{C}' \\
\quad \quad \diagup\diagdown \\
\quad \dots t \dots
\end{array}
$$
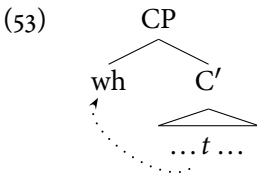
Even though, this approach may capture the extraction facts due to a lack of a higher DP shell, it suffers from an obvious and severe problem: it is unclear why the CP can occur in positions that are entirely reserved for DPs. On these grounds, this theory does not constitute an ideal approach to FRs.

5.2.3. Multidominance

In the grafting approach by van Riemsdijk (2006), the *wh*-phrase is simultaneously part of both the embedded and the matrix clause. This comes about by grafting, which is a form of multidominance. The structure is shown in (54).

(54)

$$
\begin{array}{c}
\text{VP} \\
\diagup\diagdown \\
\text{V} \quad \text{DP} \\
\quad \quad | \\
\quad \text{wh}_i \quad \dots e_i \dots \\
\quad \quad | \quad \diagdown\diagup \\
\quad \text{Spec-C} \quad \text{C}' \\
\quad \quad \diagdown\diagup \\
\quad \quad \text{CP}
\end{array}
$$

---

[24]See also Caponigro (2003).

Even though van Riemsdijk's (2006) analysis is intuitively close to the present approach in that both assume that the *wh*-phrase is part of two clauses, the grafting approach faces at least two kinds of problems: First, there are the conceptual problems related to grafting, and to multidominance more generally (it is, e.g, unclear what predictions with respect to extraction such approaches make). Furthermore, the account faces an empirical problem when it comes to case matching: it is in fact not clear how the *wh*-phrase can fulfill two conflicting case requirements as illustrated by the data discussed in section 2.2.2 and section 4.2.2.[25] In the present approach, case mismatches are in principle allowed because there are two case features, one in the embedded clause on the *wh*-phrase and one in the matrix clause on the covert head. In the grafting approach, on the other hand, there is only one case feature which cannot bear two case values at the same time.

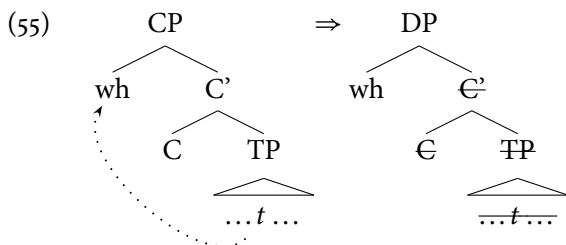### 5.2.4.   *Derived-Head Accounts*

The following summarizes and discusses two recent derived-head accounts: Ott (2011) and Donati and Cecchetto (2011).[26] In this type of theory, the *wh*-phrase is base-generated in the embedded CP and moves to a position where it becomes the head of the clause.

5.2.4.1.   *Ott (2011)*

In Ott's (2011) analysis, the FR starts out as a normal CP where the *wh*-phrase is moved to Spec-C. Then, spell-out applies not only to the complement of C but to the C head as well because it does not bear any interpretable features (in contrast to, e.g., embedded questions). Since the head of the CP has been sent to Transfer, only the *wh*-phrase remains and becomes the head of the phrase. The important steps of the derivation are shown in (55).

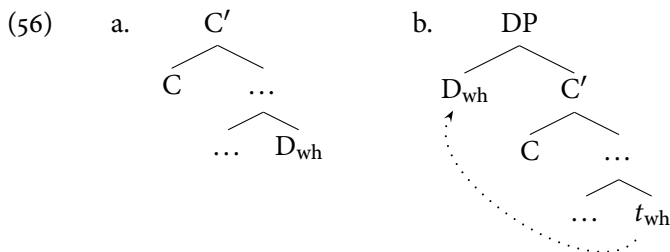[25] For this reason, van Riemsdijk (2006) argues, that there are in fact no case mismatches. See also the discussion in footnote 4.

[26] Other approaches of this kind include, e.g., Hirschbühler (1976), Bury and Neeleman (1999), Iatridou et al. (2002)

(55)

```
        CP                ⇒        DP
      /    \                     /    \
    wh      C'                 wh      C̶'̶
    ↑      /  \                      /    \
    ┊     C    TP                  C̶     T̶P̶
    ┊         / \                        / \
    ┊      …t…                        …t…
    └┄┄┄┄┄┄┘
```

Similarly to the account by Bresnan and Grimshaw (1978), Ott's account has a hard time dealing with the extraposition data, because in order to account for the facts, the DP should be able to extrapose, contrary to fact. To this end, Ott (2011: fn.5, p.186) proposes two solutions: (i) extraposition applies post-syntactically after a process of intonational phrasing that assigns a clausal intonation to the FR-DP; (ii) CPs are base-generated in the right periphery and can undergo leftward-movement. Still, both solutions do not capture the data correctly. If only constituents with a clausal intonation could undergo extraposition, PPs are not supposed to be able to extrapose, contrary to fact. The second solution does not tell us anything about the situation in (55). At the point when the FR is constructed, it is already a DP and should thus not be able to remain in the right periphery. As far as I can see, these problems cannot be overcome without referring to the syntactic labels 'DP' and 'CP'.

5.2.4.2. *Donati and Cecchetto (2011)*

Finally, there is the reprojection approach by Donati and Cecchetto (2011). The main idea here is that the *wh*-phrase is merged inside a CP and moved to Spec-C. If it is a simple *wh*-phrase like *was* ('what') it may reproject as a head and turn the CP into a DP. The derivation is sketched in (56).

(56) a.

```
        C′
      /    \
    C        …
            /  \
          …    D_wh
```

b.

```
           DP
         /     \
      D_wh       C′
       ↑        /  \
       ┊       C     …
       ┊            /  \
       ┊          …    t_wh
       └┄┄┄┄┄┄┄┄┄┄┄┘
```

Similarly to Bresnan and Grimshaw (1978) and Ott (2011), the category dominating the *wh*-phrase is a DP. Therefore, the extraposition data do not follow

straightforwardly from this analysis. Furthermore, the analysis bans complex *wh*-phrases from occuring in FRs. Donati and Cecchetto (2011) explicitly discuss this issue, claiming that FRs that contain a complex *wh*-phrase followed by *ever* are no real FRs. However, it remains unclear how, e.g., German FRs are derived where the *wh*-phrase is contained in a PP or where it is a possessor (see also Grosu 1996, 2003).

(57)   a.   Ich lade ein [FR [PP auf wen] sich auch Maria freuen   würde]
           I   invite   on who self also Maria be.happy would
           'I invite whoever Maria would also be happy to meet'
           (Vogel 2001: 904)

      b.   [FR Wessen Birne noch halbwegs in der Fassung steckt] pflegt
              whose bulb still halfway in the socket   sticks   uses
           solcherlei Erloschene zu vermeiden
           such   extinct   to avoid
           'Whoever still has half of his wits tends to avoid such vacant characters'   *(Müller 1999: 78)*

In sum, the discussion of the previous approaches has shown that all of them lack an explanation for one or the other property of FRs discussed in section 2. Most of them make empirically wrong predictions. The only empirically adequate approach by Groos and van Riemsdijk (1981) was pursued and refined above in section 3. The present version of this approach is able to give an answer to the question about the special link between the covert head and the *wh*-phrase.

6.   Conclusion

The paper has investigated free relative constructions, which pose a puzzle for syntactic theories because the *wh*-phrase that occurs in this construction seems to be simultaneously part of two sentences. While certain properties (number agreement, extraction) suggested that the *wh*-phrase is part of the matrix clause, other properties (extraposition, restricted instances of case mismatches) showed that it must be a part of the embedded clause. The ambivalent behavior of the *wh*-phrase with respect to its position is derived by distributing operations among three different components: the pre-syntactic, the

syntactic, and the post-syntactic component. Since the order of these components is strict, the derivation of FRs proceeds in three different stages.

First, a part of the *wh*-phrase, namely the part containing the argument licensing features, is copied pre-syntactically in the lexical array. Second in the syntactic component, the *wh*-phrase containing only a copy of its argument licensing features is merged inside an embedded CP, where it is *wh*-moved to Spec-C. The original argument licensing features are merged as the head of the relative clause outside the FR. Third, the copied features must be deleted post-syntactically under c-command and matching with the original features.

The present approach can be seen as a development of the theory by Groos and van Riemsdijk (1981) that assumed that the CP is headed by a covert D head. The close link between this covert head and the *wh*-phrase can now be explained: the covert head is created out of the *wh*-item; i.e., it is a part of it.

References

Agbayani, Brian (1998): Feature Attraction and Category Movement. PhD thesis, University of California, Irvine.

Agbayani, Brian and Masao Ochi (2007): Split Lexical Insertion in Parasitic Gap Constructions. Ms. California State University, Fresno and Osaka University.

Assmann, Anke, Doreen Georgi and Philipp Weisser (2012*a*): A Minimalist Analysis of Possessor Advancement. *In:* E. Boone, K. Linke and M. Schulpen, eds, *Proceedings of ConSOLE XIX (2011, Groningen)*. pp. 77–103.

Assmann, Anke (2012*b*): Deriving Parasitic Gaps by Fission and Fusion. *In:* E. Boone, K. Linke and M. Schulpen, eds, *Proceedings of ConSOLE XIX (2011, Groningen)*. pp. 49–75.

Assmann, Anke, Svetlana Edygarova, Doreen Georgi, Timo Klein and Philipp Weisser (2013): Possessor Case in Udmurt: Multiple Case Assignment Feeds Post-syntactic Fusion. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 23–63.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**(1), 35–73.

Bresnan, Joan (1973): 'Headless' Relatives. Ms., MIT.

Bresnan, Joan and Jane Grimshaw (1978): 'The Syntax of Free Relatives in English', *Linguistic Inquiry* **9**(3), 331–391.

Brosz.iewski, Ulf (2003): *Syntactic Derivations - A Nontransformational View*. Vol. 470 of *Linguistische Arbeiten*, Niemeyer, Tübingen.

Bury, Dirk and Ad Neeleman (1999): Projection of Free Relatives. Ms., University College London.

Caponigro, Ivano (2002): Free Relatives as DPs with Silent D and a CP Complement. *In:* V. Samiian, ed., *Proceedings of WECOL 2000*. California State University, Department of Linguistics, Fresno, pp. 140–150.

Caponigro, Ivano (2003): Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-linguistically. PhD thesis, University of California Los Angeles, Los Angeles, Californien.

Chomsky, Noam (1973): Conditions on Transformations. *In:* S. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

Chomsky, Noam (1980): 'On Binding', *Linguistic Inquiry* 11(1), 1–46.

Chomsky, Noam (1981): *Lectures on Government and Binding: The Pisa Lectures*. 7 edn, Mouton de Gruyter, Berlin.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* R. Martin, D. Michaels and J. Uriagereka, eds, *Step by Step. Papers in Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, chapter 3, pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. *In:* M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 1–52.

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, C. P. Otero and M. L. Zubizarreta, eds, *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud*. MIT Press, Cambridge, chapter 6, pp. 133–166.

Chomsky, Noam (2013): 'Problems of Projection', *Lingua* . available online.

Citko, Barbara (2004): 'On Headed, Headless, and Light-Headed Relatives', *Natural Language and Linguistic Theory* 22, 95–126.

Daalder, Saskia (1977): 'Over relativie zinnen met ingesloten antecedent', *Spektator* 7/8, 401–407.

Doliana, Aaron (2013): The Super-Strong Person-Case Constraint: Scarcity of Resources by Scale-Driven Impoverishment. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 177–202.

Donati, Caterina and Carlo Cecchetto (2011): 'Relabeling Heads: A Unified Account for Relativization Structures', *Linguistic Inquiry* 42(4), 519–560.

Epstein, Sam, Erich Groat, Ruriko Kawashima and Hiratsugu Kitahara (1998): *A Derivational Approach to Syntactic Relations*. Oxford University Press, New York and Oxford.

Fanselow, Gisbert (1987): *Konfigurationalität*. Narr, Tübingen.

Fox, Danny and Jon Nissenbaum (1999): Extraposition and Scope: A Case for Overt QR. *In:* S. Bird, A. Carnie, J. D. Haugen and P. Norquest, eds, *Proceedings of the*

18th West Coast Conference on Formal Linguistics. Cascadilla Press, Somerville, Massachusetts, pp. 132–144.

Fuß, Eric (2008): 'Multiple Agreement and the Representation of Inflection in the C-Domain', *Linguistische Berichte* **213**, 77–106.

Fuß, Eric and Günther Grewendorf (2012): Freie Relativsätze mit d-Pronomen. Handout. Talk given at GGS 38, Freiburg im Breisgau.

Georgi, Doreen (2012): A Local Derivation of Global Case Splits. *In:* A. Alexiadou, T. Kiss and G. Müller, eds, *Local Modelling of Nonlocal Dependencies in Syntax*. Niemeyer, pp. 305–336.

Göbbel, Edward (2007): Extraposition and PF Movement. *In:* E. Brainbridge and B. Agbayani, eds, *Proceedings of WECOL 2006*. California State University, Fresno, Fresno, California, pp. 132–145.

Gračanin-Yuksek, Martina (2008): 'Free Relatives in Croatian: An Argument for the Comp Account', *Linguistic Inquiry* **39**, 275–294.

Groos, Anneke and Henk van Riemsdijk (1981): Matching Effects with Free Relatives: A Parameter of Core Grammar. *In:* A. Belletti, L. Brandi and L. Rizzi, eds, *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*. Scuola Normale Superiore, Pisa, pp. 171–216.

Grosu, Alexander (1996): 'The Proper Analysis of "Missing-P" Free Relative Constructions', *Linguistic Inquiry* **27**(2), 257–293.

Grosu, Alexander (2003): 'A Unified Theory of 'Standard' and 'Transparent' Free Relatives', *Natural Language and Linguistic Theory* **21**(2), 247–331.

Grosu, Alexander and Fred Landman (1998): 'Strange Relatives of the Third Kind', *Natural Language Semantics* **6**, 125–170.

Harbert, Wayne (1983): 'On the Nature of the Matching Parameter', *The Linguistic Review* **2**, 237–284.

Heck, Fabian (2010): Categories, Recursion and Bare Phrase Structure. Talk at GGS, Freie Universität Berlin.

Heck, Fabian and Malte Zimmermann (2004): Remarks on the Phasal Status of DP. Ms., Universität Tübingen, Humboldt-Universität Berlin.

Hirschbühler, Paul (1976): Two Analyses of Free Relatives in French. *In:* A. Ford, J. Reighard and R. Singh, eds, *North East Linguistic Society (NELS) 6*. Montreal Working Papers in Linguistics, Montreal, pp. 137–152.

Hirschbühler, Paul (1978): The Semantics and Syntax of wh-Constructions. PhD thesis, University of Massachusetts, Amherst.

Hirschbühler, Paul and María-Luisa Rivero (1981): 'Catalan Restrictive Relatives: Core and Periphery', *Language* **57**, 591–625.

Iatridou, Sabine, Elena Anagnostopoulou and Roumyana Izvorski (2002): Some Observations about the Form and Meaning of the Perfect. *In:* M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 189–238.

Inaba, Jiro (2005): Extraposition and the Directionality of Movement. *In:* S. Blaho, L. Vicente and M. de Vos, eds, *Proceedings of ConSOLE XIII*. University of Leiden, Leiden, pp. 157–169.

Jacobson, Pauline (1995): On the Quantificational Force of English Free Relatives. *In:* E. Bach, E. Jelinek, A. Kratzer and B. H. Partee, eds, *Quantification in Natural Languages*. Kluwer, Dordrecht, pp. 451–486.

Keine, Stefan (2010): *Case and Agreement from Fringe to Core: A Minimalist Approach*. Vol. 536 of *Linguistische Arbeiten*, De Gruyter, Berlin.

Kiparsky, Paul (1971): Historical Linguistics. *In:* W. O. Dingwall, ed., *A Survey of Linguistic Science*. University of Maryland Linguistics Program, College Park, pp. 576–642.

Kiparsky, Paul (1976): Abstractness, Opacity, and Global Rules. *In:* A. Koutsoudas, ed., *The Ordering and Application of Rules*. Mouton.

Kuroda, Sige-Yuki (1968): 'English Relativization and Certain Related Problems', *Language* **44**, 244–266.

Larson, Richard K. (1987): '"Missing Prepositions" and the Analysis of English Free Relative Clauses', *Linguistic Inquiry* **18**(2), 239–266.

Link, Godehard (1983): The Logical Analysis of Plural and Mass Terms: A Lattice-Theoretical Approach. *In:* R. Bäuerle, C. Schwarze and A. von Stechow, eds, *Meaning, Use and Interpretation of Language*. de Gruyter, Berlin, pp. 302–323.

McCawley, James (1970): 'English as a VSO Language', *Language* **46**, 286–299.

Müller, Gereon (2011): *Constraints on Displacement: A Phase-Based Approach*. Vol. 7 of *Language Faculty and Beyond*, John Benjamins, Amsterdam.

Müller, Gereon and Wolfgang Sternefeld (1993): 'Improper Movement and Unambiguous Binding', *Linguistic Inquiry* **24**, 461–507.

Müller, Stefan (1999): 'An HPSG-Analysis for Free Relatives in German', *Grammars* **2**, 53–105.

Nunes, Jairo (1995): The Copy Theory of Movement and Linearization of Chains in the Minimalist Program. PhD thesis, University of Maryland, College Park.

Nunes, Jairo (2004): *Linearization of Chains and Sideward Movement.*. MIT Press, Cambridge, Massachusetts.

Ott, Dennis (2011): 'A Note on Free Relative Clauses in the Theory of Phases', *Linguistic Inquiry* **42**(1), 183–192.

Pittner, Karin (1991): Freie Relativsätze und die Kasushierarchie. *In:* E. Feldbusch, ed., *Neue Fragen der Linguistik Vol. 1*. Niemeyer, Tübingen, pp. 341–347.

Pittner, Karin (1995): 'Regeln für die Bildung von freien Relativsätzen', *Deutsch als Fremdsprache* **32**(4), 195–200.

Postal, Paul (1972): 'On Some Rules That Are Not Successive Cyclic', *Linguistic Inquiry* **3**, 211–222.

Pullum, Geoffrey K. (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York.

Riemsdijk, Henk van (2006): Free Relatives. *In:* M. Everaert and H. v. Riemsdijk, eds, *The Blackwell Companion to Syntax*. Blackwell, Oxford, pp. 338–382.

Rooryck, Johan (1994): 'Generalized Transformations and the wh-Cycle: Free Relatives as Bare wh-CPs', *Groninger Arbeiten zur germanistischen Linguistik* 37, 195–208.

Rouveret, Alain and Jean-Roger Vergnaud (1980): 'Specifying Reference to the Subject: French Causatives and Conditions on Representations', *Linguistic Inquiry* 11(1), 97–202.

Suñer, Margarita (1984): 'Free Relatives and the Matching Parameter', *The Linguistic Review* 3, 363–387.

Svenonius, Peter (2004): On the Edge. *In:* D. Adger, C. de Cat and G. Tsoulas, eds, *Peripheries: Syntactic Edges and their Effects*. Kluwer, Dordrecht, pp. 261–287.

Thomas, Daniela (2013): Split Ergativity in Subordination. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 1–21.

Trommer, Jochen (2006): Plural Insertion is Constructed Plural. *In:* G. Müller and J. Trommer, eds, *Subanalysis of Argument Encoding in Distributed Morphology. Volume 84 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 197–228.

Trommer, Jochen (2008): A Feature-Geometric Approach to Amharic Verb Classes. *In:* A. Bachrach and A. Nevins, eds, *Inflectional Identity*. Oxford University Press, pp. 206–236.

Truckenbrodt, Hubert (1995): Extraposition from NP and Prosodic Structure. *In:* J. Beckman, ed., *Proceedings of the North Eastern Linguistics Society 25*. GLSA Publications, Amherst, Massachusetts, pp. 503–517.

Vogel, Ralf (2001): Case Conflict in German Free-Relative Constructions: An Optimality-Theoretic Treatment. *In:* G. Müller and W. Sternefeld, eds, *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 341–375.

Wurmbrand, Susi (2007): 'How Complex Are Complex Predicates?', *Syntax* 10(3), 243–288.