IM NOT PERFECT
The case against copying

Jan Koster

1. Introduction and theoretical background

One of the major issues in syntactic theory has been the idea of successive cyclic Wh-movement. After its introduction in Chomsky (1973), the elimination of unbounded movement was a hotly debated topic and in the middle of these discussions, it was observed by du Plessis (1977: 725) that languages like Afrikaans seem to show doubling of Wh-phrases as in (1):

(1)   *Waarvoor* dink julle *waarvoor* werk ons?   Afrikaans
         wherefor think you wherefor work we
         'What do you think we are working for?'

Not much later, similar observations were made for Frisian and numerous other languages (see Felser 2004 and particularly Den Dikken 2009 for much discussion and references). Hiemstra (1986: 99), another early source of such data, gives examples like the following (with, interestingly, the non-interrogative complementizer cliticized on the second occurrence of *wêr*):

(2)   *Wêr*   tinke jo   *wêr*'t   Jan wennet?   Frisian
         where think you where that-CL Jan resides
         'Where do you think that John lives?'

At the time, such examples were widely interpreted as showing successive cyclic movement in action: the second and further copies were seen as the footprints left behind by Wh-phrases moving from COMP to COMP.
        Particularly since Chomsky (1993), the data in question have come into the limelight again, this time interpreted as evidence for the so-called copying theory of movement (nowadays seen as Internal Merge, henceforth IM) (see, for instance, Hornstein, Nunes and Grohmann 2005: 215). This interpretation gets extra weight from the fact that Chomsky has argued that copying is not just a possible analysis but a sign of the perfection of the design of human language (see, for instance, Chomsky 2007: 8-11).
        In this article, I will argue that the ubiquitous copying analysis for data like (1) and (2) is mistaken and that, more generally, IM and copying are not signs of perfection, but operations excluded by a reasonable principle of Universal Grammar (or a different standard of perfection, if you will).
        This can be seen as follows. Assuming that syntactic structures show hierarchical structure and recursion -the long-standing consensus-, we observe that such structures are used to define two patterns of secondary computation: completions and additions. At this point, I will limit myself to completions and I will briefly discuss additions in the concluding section. A typical example of a completion is the interpretation of anaphors like *herself*:

(3)   *Mary* saw *herself*

The anaphor *herself* can, referentially, not stand on its own feet and is therefore completed by its connection to the antecedent *Mary*. This is a dependency relation between a dependent element D and an antecedent A and all such relations have the properties of what I called "the Configurational Matrix" (henceforth CM, see Koster 1987):[1]

(4)   **Configurational Matrix**
         A completion (a dependency relation A-D) is:

         a.   obligatory
         b.   bi-local
         c.   binary
         (d.   asymmetrical: A precedes D)

All completions are **obligatory**: a sentence like (3) would be ungrammatical without an appropriate antecedent for *herself*. The relation is **bi-local**, i.e. there is a locality condition on A (also known as c-command and nowadays said to involve a minimal search domain for probe-goal interaction) and there is a locality condition on D (some local domain B in which D has to find an antecedent A). Relations of type A-D are **binary** in that there is exactly one A for D (no "split" antecedents) and one D for A.

         The fourth property is about linear order, particularly about the fact that usually A precedes D rather than the other way around. I have put it in parentheses in accordance with the general feeling (shared by me) that linearization is not a "deep" core property of grammar but something having to do with the usability of language. Nevertheless, I believe that precedence of A over D is a universal characteristic of completions, in accordance with what I see as a natural principle of word order (number (1), as another one will be mentioned at the end of this article):

(5)   **Principle of Natural Word Order (1)**
         It is natural for a dependent element (D) to follow the source (A) of its completion

Potentially, this is more than a stipulation and I will come back to it in section 5. It is a typical "third factor" principle in the sense of Chomsky (2007) in that it seems plausible that perceptually optimal computation involves the order A-D rather than D-A. With the order A-D, resolution is instantaneous when D appears, as A is already given. With the order D-A, resolution of D has to wait till A appears, thereby causing extra memory load. I see, incidentally, (5) as part of an alternative to the LCA of Kayne (1994), a matter I will leave at rest here (but see Koster 2008).

         In earlier work (Koster 1987: 8), I characterized the completion relation between A and D as being governed by an interaction called **"share property."** At the time, "share property" was presented as an alternative to "move alpha." The latter was rejected because "move alpha" was much less general than "share property" in that "move alpha" was used exclusively for completions involving gaps. This arbitrary restriction on the mechanism of local feature exchange (*in casu* completion) was not motivated, neither in the 1980s nor in its current, allegedly perfect avatar as IM-cum-copying.

         Unlike "move alpha," the more general mechanism "share property" entails partial reconstruction, namely of those features that are missing in D, under the further assumption

---

[1] Note that the CM is a meta-theoretical statement and, as such, neutral between derivational and representational approaches to syntax.

that each D has quite a few features of its own. Since, in other words, each D has features of its own, not all of its features have to be reconstructed via A. "Move alpha" (or IM-cum-copying) in contrast, involves total reconstruction, which I see as a residue of early generative grammar with its Harrisian "normalization" philosophy, which motivated total-reconstruction levels like deep structure. In my opinion, it became obvious during the 1970s that total reconstruction does not exist. Since the late 1960s, one after another form of total reconstruction was rejected and for good reasons (see for instance Wasow (1972) for discussion of the inadequacy of earlier rules involving total reconstruction, like Pronominalization, Equi-NP Deletion and many others).

This very promising shift from total reconstruction to partial reconstruction culminated in the landmark dissertation of Roger Higgins (1973). Analyzing pseudocleft sentences, Higgins concluded that total reconstruction is impossible in these sentences and that in a sentence like *What John saw was himself* there is no way to relate *himself* to the object position of *saw* by movement transformations without violating the usual constraints, like island conditions and even strict cyclicity (see also Blom & Daalder 1977). Nevertheless, *himself* behaves as if it is c-commanded by *John*, the so-called correspondence problem. This problem is still with us, and it shows that a lexical value like *himself* can be separated from the position in which it is interpreted without movement but, instead, by the mediation of a pronoun and a predication specifying the value of that pronoun. Similar considerations apply to left-dislocated phrases in Dutch and other languages (see Koster 1987: 32ff.) and from such observations I concluded that the concept of "movement," in all its forms, is the main obstacle for a unified theory of linguistic dependencies. It still is, in my opinion, in its current reincarnation as Internal Merge (IM). Rather than a sign of progress or perfection of design, IM is symbolic of the tenacity with which the total reconstruction paradigm is maintained.

Another way of looking at what is wrong with IM and copying theories is the assumption that UG includes the following principle:

(6)   **The Anaphoric Principle**
         A relation A-D (i.e., any relation with the properties of the CM) is possible if and only if D is less complete (feature-wise) than A

I consider this a well-corroborated core principle of UG. The minimalistically inclined can see (6) as an economy or "least effort" principle: feature exchange only takes place if necessary (i.e., if triggered by incompletion). Principle (6) excludes Principle C of the classical Binding Theory (about disjoint reference (DR) of R-expressions) as a completion relation A-D, which is confirmed by the fact that it does not show the properties of the CM (6): it does not even require locality or c-command (see Koster 1987: 346ff. for discussion).

Most important in the current context, the Anaphoric Principle also excludes copying theories, as two copies are equally complete by definition. Rather than being an instantiation of perfect design, in other words, copying is excluded by a well-corroborated principle of UG. Luckily, there are alternatives for the movement theories of Wh-displacement, as briefly considered by Chomsky (1973: 284) and erroneously dismissed on the same page (Chomsky's

footnote 70).[2] According to this alternative, both Wh-phrases and the corresponding gaps are generated *in situ* (by External Merge, if you prefer) and exchange the necessary features by "share property." This is the only account of Wh-displacement to date that makes possible unification of A-D relations in terms of the Configurational Matrix and the Anaphoric Principle.

It also is the only account of Wh-gap relations that does justice to the fact that reconstruction is not total but partial: the gaps have properties independent of the dislocated Wh-phrases, like theta-roles and Case. Moreover, gaps have different properties when locally bound (by any XP category) than when they are bound across islands. In the latter case, the gaps are usually limited to the category DP, as was first observed by Cinque (1983) and Obenauer (1984) (see Koster 1987: 153ff. for discussion). Gaps corresponding to Wh-phrases behave like anaphors in certain respects in local contexts and like empty resumptive pronouns across islands (like parasitic gaps and other relatively acceptable cases; see also Cinque 1990, Postal 1993 and Ouhalla 2001).[3] The distinction anaphoric-pronominal, which makes only sense for empty elements and not for full reconstruction copies, which are the same in both local and non-local contexts. In fact, copies, being equally complete, have no binding potential whatsoever, let alone the contextually differentiated potential observed for gaps. I take the Cinque-Obenauer observation therefore as strong *prima facie* evidence against a copying -full reconstruction- theory of "Wh-movement."

Before discussing some actually existing copying analyses, it is important to emphasize that the Anaphoric Principle expresses the generalization that (feature-wise) the higher (preceding) link A is always more specified than D. Thus, in a sentence like *Mary saw herself*, *Mary* is the "richer" element, providing both referential identity and further identity features to *herself*.[4] In accordance with the Principle of Natural Word Order, most information flows from left to right, and, since the relation involves c-command, from higher to lower. Exactly what we see with anaphora, can be observed for Wh-gap constructions, as is to be expected:

(7)   *Which man* did you see [*e*]

---

[2] The problems raised in the footnote involve classical cases that are supposed to show the need for full reconstruction, like *Which pictures of each other were the men looking at [e]*. This is precisely the kind of example that has no force if we appreciate the consequences of Higgins' (1973) insights about pseudocleft constructions. Thanks to the anaphoric properties of the trace ([*e*]), the lexical value [*which pictures of each other*] is available through "share property," annihilating the need for full reconstuction. The other problem raised is a phrase like the *headway that he made.* If one follows the logic of full reconstruction, the so-called raising analysis for relative clauses (Brame, Vergnaud 1974 , and more recently Kayne 1994) is practically necessary. However, I hope to show elsewhere that the raising analyis does not work because it violates island conditions when the relative pronous pied-pipe a larger phrase. So, in fact the relative clause example is not evidence in favor of full reconstruction but evidence against it. Next to pseudoclefts and left-dislocations, it is another example of partial reconstruction (in this case mediated by the relative pronoun).

[3] By "anaphoric" I mean "bound in a local domain," not close resemblance to, say, reflexive binding in all other respects.

[4] The information flow from A to D, in that case, is mainly from left to right. However, this is due to (5), not to the relation "share property" itself, which is symmetrical. This can be concluded from a sentence like *008 shaved herself.* You conclude from that that 008 is a female, even if you have never heard of agent 008 before.

In these structures, the antecedent (*which man*) is typically richer in feature content than the gap, which only has subcategorization, Case and theta-role features. The identity of [e] is completed by the fact it shares the features of the more specified antecedent *which man*. This order of information flow is overwhelmingly true for all chains supposed to result from movement, as it is for all cases classified as NP-movement, Wh-movement and non – morphological V-movement (like Verb Second in Germanic). This leads to the *next prima facie* argument against chain formation by copying.

Barbiers, Koeneman and Lekakou (2008), inadvertently, point out an important problem for copying theories about Dutch, namely the fact that the alleged copies are not always the same. With some regional variation, copying phenomena are quite common in Dutch.[5] A standard case (also grammatical to my ear) is the following:

(8)   *Wie* denk je   *wie* ik gezien heb?
        who think you who I   seen have
        'Who do you think I have seen?'

The two copies (*wie*) are exactly the same. However, as Barbiers *et al.* point out, there are also three regional variants found with non-equal copies:

(9)   *Wat* denk je   *wie* ik gezien heb?
        what think you who I   seen have

(10)   *Wie* denk je *die*   ik gezien heb?
        who think you rel.pron I   seen have

(11)   *Wat* denk je *die*   ik gezien heb?
        what think you rel.pron I seen have

Barbiers *et al.* claim that the italicized pairs of items in each sentence form chains of copies. Since the alleged copies can be rather different, as shown by (9-11), the first thought that comes to mind is that these examples are counterexamples to the copying theory. However, instead Barbiers *et al.* protect the copying theory by the *ad hoc* notion of "partial copying," thereby implicitly giving up the intuition that motivated the copying theory in the first place, namely the normalization intuition of full reconstruction.

Partial copying cannot be implemented without further *ad hoc* procedures, such as remerging features from below the word level, which violates all known island conditions. In order to do that, Barbiers *et al.* postulate a layered nominal projection, with the features of *what* at the lowest level, *wie* at the intermediate level, and *die* at the highest level. This is supposed to correspond with a specificity hierarchy, with *wat* as least specific element and *die*

---

[5] It is far from clear whether doubling phenomena in Dutch correspond with well-defined geographical micro-variation. Although I am a speaker of standard Dutch, my intuitions about the phenomena in question completely agree with what is reported by the dialect speakers whom I have consulted. This is a situation reminiscent of parasitic gaps: most people do not realize that they are in their grammar until somebody tells them so, in which case judgments tend to converge. Nevertheless, there is obvious geographical and regional variation in the acceptance of doubling phenomena. Just to be sure, I have checked all my judgments with informants who can be considered stable speakers of a doubling dialect. Particularly, I would like to thank Jakolien den Hollander, who is a speaker of an Overijssel variety of Dutch in which doubling and question formation with clause-initial *wat* is considered normal. Any remaining errors are my own.

as most specific element, and *wie* in between. This classification is not unattractive, but it is a far cry from a demonstration that subparts of the postulated projection can move (viz. be copied). I conclude therefore that the notion of partial copying is *ad hoc* and in fact undermines the notion of copying altogether.

However, things become even worse as Barbiers *et al.* claim that their pairs of copies form links in a movement chain (what copying is all about) with the following property derived from their specificity hierarchy (*op. cit.*, p. 78):

(12)   In a syntactic movement chain, a higher chain link cannot be more specified than a lower chain link

This is supposed to be shown by the fact that the "copies" in (9-11) do not occur in reverse order (similar facts are discussed by Den Dikken and Bennis 2009):

(13)   *__Wie__ denk je   __wat__ ik gezien heb?
          who think you what I   seen have

(14)   *__Die__   denk je __wie__ ik gezien heb?
          rel.pron think you who I   seen have

(15)   *__Die__   denk je   __wat__ ik gezien heb?
          rel.pron think you what I seen have

The generalization (12) derived from (13-15) is completely at variance with what has always been obvious about chain links and just the opposite from the universal expressed by the Anaphoric Principle (6). Since I agree with Barbiers *et al.* that sets of "copies" are often in agreement with (12) and given the fact that they violate a very well-corroborated universal like the Anaphoric Principle, it is a near-certainty that the associated Wh-words in so called copying constructions (like 9-11) are in fact not chain links of movement or copies at all. In short, the partial copying facts of Barbiers *et al.* are very interesting but show exactly the opposite from what the authors claim that they show. So, if multiple, associated occurrences of Wh-words are not copies, what are they?

2. Problems of copying theories

First, I would like to expand the set of data somewhat as to also include facts that were sometimes discussed under the heading "partial Wh-movement" (see Van Riemsdijk 1983 and McDaniel 1989, among others). The following data are from Van Riemsdijk (1983):

(16)
    a. **Was** glaubst du, **was** Peter meint,   **was** Hans sagt, **was** Klaus behauptet, **mit**   **wem**$_1$
       What think you what Peter believes what Hans says what Klaus claims   with   whom
       Mary t_1 gesprochen hat
       Maria   spoken   has
    b. **Was** glaubst du, **was** Peter meint, **was** Hans sagt, **mit wem** Klaus behauptet, dass Maria
       gesprochen hat
    c. **Was** glaubst du, **was** Peter meint, **mit wem** Hans sagt, dass Klaus behauptet, dass Maria
       gesprochen hat
    d. **Was** glaubst du, **mit wem** Peter meint, dass Hans sagt, dass Klaus behauptet, dass Maria
       gesprochen hat
    e. **Mit wem** glaubst du, dass Peter meint, dass Hans sagt, dass Klaus behauptet, dass Maria
       gesprochen hat

These sentences, roughly, are variants of each other, with the Wh-phrase *mit wem* having scope over the whole sentence. Variant (16e) has long Wh-movement like English. However, as shown by (16a-e), the Wh-phrase can show up at the beginning of any of the intermediate clauses, in which case the scope is computed via repeated occurrences of *was*, all the way to the highest clause, where the scope is ultimately fixed. Originally, *was* was seen as nothing other than a scope marker (as in Van Riemsdijk 1983 and McDaniel 1989 for Romani), with the movement of the "real" Wh-phrase (*mit wem*) described as "partial movement." Gradually, however, other approaches came to the fore, according to which *was* is not seen as just a scope marker but as a regular Wh-phrase corresponding to an argument DP (see the valuable collection of articles edited by Lutz, Müller and Von Stechow 2000). The original approach is now often referred to as the Direct Dependency Approach, while the other approach, which goes back to Dayal 1994, is called the Indirect Dependency Approach.

Although in earlier work, I adopted the Direct Dependency approach (see Koster 1987 and 2003), I now believe that a variant of the Indirect Dependency Approach is correct. The approach I will assume involves "concealed relatives," in Dalina Kallulli's felicitous phrase (2005) and comes close to a proposal made by Felser (2001), be it in a rather different implementation. It is my claim that a solution based on concealed relatives is not only the solution to the proper analysis of the "partial movement" facts as in (16), but also the key to understanding what so-called copies of Wh-phrases really are. The solution I have in mind not only seems to solve the occurrence of what looks like real copies (like (8) above), but also the possibility of partial and other deviant forms of copying.

Before giving my solution, I would like to raise the suspense some more by giving a short summary of problems not solved by any copying theory so far. The first three problems (17a-c) are well-known and derived here from Schippers (2009). The fourth problem is my own addition, discussion of which will be the *pièce de résistance* of this article, in the sense that I believe it crucially favors my own account over copying theories:

(17)   a.   Wh-phrases show up in positions where they do not seem licensed
        b.   Wh-phrases are only spelled out in intermediate SpecCPs
        c.   Copying of complex Wh-phrases is prohibited
        d.   Copies are often partial and unequal, with a pattern unexpected
                and unexplained under copying theory

2.1 Unlicensed positions
I will now discuss these problems in turn and show in the next section that they basically follow from an alternative to the copying theory. As a reference language, I use German,

which shows all three patterns of question formation under discussion here (examples cited from Schippers 2009):

(18)   a.   [$_{CP1}$ *Was* meinst du [$_{CP2}$ *wen* Marie geküsst hat?]]
                        what think you   who Marie kissed has

       b.   [$_{CP1}$ *Wen* meinst du [$_{CP2}$ *wen* Marie geküsst hat?]]
                        who think   you   who Marie kissed has

       c.   [$_{CP1}$ *Wen* meinst du [$_{CP2}$ dass Marie geküsst hat?]]
                        who think   you   that Marie kissed has
                 All three: 'Who do you think Marie has kissed'

As can be observed in both (18a) and the alleged copying structure (18b), the question word *wen* can show up in these constructions in a position selected by *meinen* ('to think'). Normally, the Wh-phrase of a question only occurs at the beginning of a sentence or in a position selected by a verb like *fragen* ('to ask') and not in the complement of a verb like *meinen*, which lacks the relevant features. Cf.:

(19)   a.   Er fragt [*wen*   Marie geküsst hat]
                 he asks who Marie kissed has

       b.   *Er meint [*wen*   Marie geküsst hat]
                 he thinks who Marie kissed has

This is a serious problem for the copying theory and also for the "partial movement" analysis of (18a), especially when one believes that there is such a thing as Wh-movement driven by feature checking. If Wh-movement in CP2 of (18a) and (18b) is not driven by Wh-features, what other feature could attract the Wh-phrase? An EPP-feature has been suggested in the literature, but that's the name of a problem rather than a solution and it would not work for (19b) anyway. Schippers (2009) cites Stepanov and Stateva (2006), who claim that heads of intermediate CPs contain a Q-feature to be checked as well, thus restating the problem rather than solving it. Moreover, they leave familiar assumptions intact, like the idea that *was* in partial movement constructions is a scope marker and, worst of all, that intermediate Wh-phrases have anything to do with questions. All in all, I think, it is fair to say that the problem of Wh-movement of question-phrases to positions not marked for corresponding question features has not been solved.

Eventually, the fact that a non-question taking verb like *meinen* selects a Wh-phrase in examples like (18a) and (b) points in a rather different direction, namely that the selected Wh-phrase does not introduce a Wh-question at all but a different type of Wh-headed clause. As I already have hinted at, to be further motivated in what follows, the alternative clause type turns out to be a relative clause.

2.2 Wh-phrases only in intermediate SpecCPs
The second problem discussed by Schippers is the fact that scope markers like *was* do not license a Wh-phrase *in situ*. Thus, the following fact is unexpected:

(20)   *[$_{CP1}$ *Was* meinst du [$_{CP2}$ Marie hat *wen* geküsst?]]
                 what think you   Marie has who kissed

Normally, Wh-phrases that introduce Wh-questions license Wh-*in situ*, even if the unmoved Wh-phrases are in a lower clause (or even in an island):

(21)   *Who* said [that Mary read *what*]?

Copying (doubling) is also impossible in configurations like (20):

(22)   *Wen meinst du [Marie hat *wen* geküsst?]
          who think   you Marie has who kissed

This example highlights a general problem for copying theories, namely that the alleged copies only show up at the edges of clauses, in the Specs of CPs. This is unexpected because copying (as a realization of Internal Merge) is supposed to apply to all former Wh-phrase-trace combinations, not just to Wh-phrases in the Specs of CPs. Under the alternative, according to which so-called copies introduce relative clauses, facts such as (20) and (22) are immediately explained: relative clause Wh-phrases do not occur *in situ*.
         Interestingly, if the copies are in SpecCP, they allow Wh-*in situ* elsewhere in the same clause, as was observed by Dayal (1994: 140):

(23)   *Was* glaubst du *wann* Hans *an welcher Universität* studiert hat?
         what think you when H. at which university studied has
         'When do you think Hans studied at which university?'

This fact is unexpected if the *in situ* phrase (*an welcher Universität*) is in a clause not headed by a question Wh-phrase but by the Wh-phrase of a relative clause (*wann* in this example). However, I will show later on that this is generally possible under certain conditions.

2.3 No copying of complex Wh-phrases
The third problem discussed by Schippers (2009) is the fact (if it is a fact) that copying usually involves simple Wh-phrases, like the German equivalent of *what*, *who*, and *with whom* at best. Complex Wh-phrase with a nominal head, like the English equivalent of *which books* are generally excluded:

(24)   *[CP1 *Welche Bücher* glaubst du [CP2 *welche Bücher* sie gekauft hat?]]
          which books believe you   which books   she bought has?

Interestingly, complex Wh-phrases are possible with the other construction, the one which is supposed to have *was* as a scope marker:

(25)   [CP1 *Was* glaubst du [CP2 *welche Bücher* sie gekauft hat?]]
          what believe you   which   books   she bought has

Dutch data, to the extent that people speak the idiolect in which the sentences in question are acceptable, are similar (although many find the Dutch equivalent of (24) somewhat better than indicated). As for other Dutch data, Schippers did extensive corpus research, which shows that combinations with so-called R-words (Van Riemsdijk 1978) are particularly frequent:

(26)   *Waar* denk je *waarop*   dat gebaseerd is?
          where think you whereon that   based is
          'On what do you think that was based'

This type of doubling is so common that it is doubtful if its acceptance is based on any well-defined dialect boundaries. To my ear, at least, examples like (26) are entirely natural. Interestingly, given the relative unacceptability of (24), Schippers also found examples like:

(27)   *Welke* denk je *welke* sneller aan het eind van zijn Latijn is en waarom?
          which think you which faster at the end of his Latin is and why
          'Which one do you think is faster at the end of his Latin and why'
          Idiomatic for: 'Which one do you think is exhausted faster and why?'

3.4 Partial and unequal copies
The fourth and, I believe, most insurmountable problem for any copying theory is the fact that, more often than not, the Wh-phrases at the beginning of the sentence differ from the following "copies." Moreover, very often the various Wh-phrases that differ from one another can only appear in one order and not in reverse order. This was in fact what Barbiers et al. observed and what led them to the idea of partial copying. I repeat (9) - (11) here for convenience:

(9)   **Wat** denk je *wie* ik gezien heb?
          what think you who I   seen have

(10)   **Wie** denk je ***die***   ik gezien heb?
          who think you rel.pron I   seen have

(11)   **Wat** denk je ***die***   ik gezien heb?
          what think you rel.pron I seen have

It is important to bear in mind that the second copy in (10) and (11) is generally seen as a relative pronoun while the first and leftmost copy is a question word. In other words, on the basis of these sentences alone, the notion "copy" is a misnomer. As we further saw, the opposite orders are impossible ((13) – (14), repeated here):

(13)   ***Wie** denk je **wat** ik gezien heb?
           who think you what I   seen have

(14)   ****Die***   denk je **wie** ik gezien heb?
           rel.pron think you who I   seen have

(15)   ****Die***   denk je **wat** ik gezien heb?
           rel.pron think you what I seen have

At least for (14) and (15), we do not need to have recourse to *ad hoc* hypotheses like partial copying supplemented with a dubious stipulation about chains. As I will demonstrate, unlike the grammatical (10) and (11), (14) and (15) are out because they do not contain the interrogative Wh-phrase necessary for their interpretation as questions.

There are examples that are superficially compatible with partial copying, like the very common doublings like (26) (repeated here):

(26) *Waar* denk je *waarop* dat gebaseerd is?
      where think you whereon that   based   is
      'On what do you think that was based'

This example, at least, would be compatible with partical copying in that *waar* can be extracted from the PP *waarop* without violating island constraints (in a somewhat different sentence):

(28) *Waar*$_i$ denk je dat dat [$_{PP}$ [*e*]$_i$ *op*] gebaseerd is?
      where think you that that   on   based   is
      'What do you think that that is based on?'

However, a closer look will reveal that (26) does not show any copying at all and that, therefore, no partial copying is involved. An immediate counterexample to partial copying approaches to examples like (26) is the kind of doubling mentioned by Den Dikken (2009, (45)) (and taken from Fanselow & Ćavar 2001):

(29) *Wieviel* sagst du *wieviel* Schweine ihr   habt?
      how.many say you how.many pigs   you.PL have
      'How many pigs do you say that you have'

The first copy, *wieviel* is a subpart of the second copy, *wieviel Schweine*. However, in this case we have a left-branch prohibition on extraction, which makes the phrase *wieviel Schweine* an island for Wh-movement:

(30) **Wieviel*   sagst du [[*e*] *Schweine*] ihr   habt
       how.many   say you   pigs   you.PL have

The same is true for Dutch. Thus, the following is generally accepted by speakers whose dialect allows doubling:

(31) *Hóéveel*   zeg je *hoeveel varkens* je   gezien hebt?
      how.many say you how.many pigs   you.PL seen have
      'How many pigs do you say (that) you saw?'

Like in German, some Left Branch Condition prohibits extraction by Internal Merge seen as partial copying:

(32) **Hoeveel*   zeg je [[e]   *varkens*] je   gezien hebt?
       how.many say you   how.many pigs   you.PL seen have
       'How many pigs dis you say (that) you saw?'

Strong counterevidence to the idea of partial movement comes from the following interesting minimal pair in Dutch:

(33) a. ?*Met wie* zei je *met wie*   je gesproken had?
          with whom said you with whom you talked   had
          'With whom did you say that you talked?'

      b. *Wie*   zei je *met wie*   je gesproken had?
          who(m) said you with whom you talked   had

Example (33a) shows the doubling of PPs that was originally observed for Afrikaans by du Plessis (1977). Most speakers of the relevant variety of Dutch accept (33a) but they all agree that (33b) is better. This is totally unexpected under a copying theory and it forms a strong counterexample against the idea of partial copying as well. The reason is that it has been known at least since Van Riemsdijk (1978) that PPs are strong islands in Dutch, with the exception of postpositional structures like PPs with R-pronouns. Thus the following is completely ungrammatical in all varieties of Dutch:

(34)   *Wie_i zei je [[$_{PP}$ *met* [e]$_i$] je gesproken had]?
              who said you   with   you   talked had

Eventually, I will claim that the other construction, with initial German *was* (*wat* in Dutch), is also an example of non-matching copies, or rather that, like what first looked as full copying, is no copying at all, but cases in which an initial interrogative phrase is specified by free relative clauses.
        In sum, then, we can conclude that, even on the limited evidence reviewed so far, partial copying is not likely to exist in Dutch and German and that, as a consequence, all cases of non-matching copies form counterevidence to the very notion of copying. I will now present my alternative analysis.

3. An alternative to copying

As I have hinted at several times already, my solution to the (pseudo)copying problem is that of chains of copies, only the first link is an interrogative Wh-phrase. The Wh-phrases introducing subsequent clauses are not copies but the Wh-phrases introducing free relatives. Consider a Dutch cleft sentence (35a) and a corresponding pseudocleft sentence (35b):

(35) a. Het is een boek *dat   hij leest*
          it   is a   book that he reads

      b. *Wat   hij leest* is een boek
          what he reads is   a   book

Both constructions involve a free relative clause. As can be seen in (35), these relative clauses come in two varieties: in (35a) the relative clause begins with a d-word (*dat*) and in (35b) it begins with a w-word (*wat*). There has been considerable debate about the exact analysis of (pseudo)cleft sentences, turning around questions as to what is the subject (with subject reversal or not), what is the nature of the predicate (specifying or equative) and so on (see Higgins 1973, Blom en Daalder 1977, Moro 1990, Williams 1994, Heycock & Kroch 1999, Den Dikken 2006 and many others). I will not go into these controversies but simply assume

that these sentences are specificational and involve two different forms of the free relative clause.

The two variants of relative clauses shown in (35a) and (35b) are as follows. The pseudocleft sentence (35b) involves an implicit head (called an included antecedent in the Dutch grammatical tradition). Unlike what we see in English, this implicit antecedent can be spelled out in Dutch, with (36) as a variant of (35b):

(36)   *Dat   wat*   hij leest is een boek
       that what he reads is a   book

In the cleft variant (35a), there is no implicit head that can be spelled out:

(37)   *Het is een boek *dat   wat* hij leest
        het is   a   book that what he reads

Otherwise, cleft sentences show the normal range of relative pronouns of Dutch, including obligatory w-words with prepositions:

(38)   Het is de man *met wie*   hij gesproken heeft
       it   is the man with whom he talked   has
       'It is the man with whom he talked'

There is a simple generalization about when a spelled-out head of a relative clause is prohibited in Dutch, namely when the relative clause specifies the expletive *het* (as in 35a). The same can be observed with object expletives, as with verbs like *regret*. Thus, there are two possibilities for free relatives in such cases, dramatically distinguished by occurrence on different sides of the verb in Dutch:

(39)   a.   Hij heeft *het* betreurd *wat   Peter doet*
              he has   it regretted what Peter does

        b.   Hij heeft *wat Peter doet* betreurd
              he has   what Peter does regretted
              (both: 'He has regretted what Peter does')

As predicted, an implicit head can only be spelled out in (40b), corresponding to (39b):[6]

(40)   a.   *Hij heeft het betreurd *dat   wat   Peter doet*
              he has   it regretted that what Peter does

        b.   Hij heeft *dat wat   Peter doet* betreurd
              he has   that what Peter does regretted

More generally, these facts are in accordance with a generalization captured by Stowell's Case Resistance Principle, namely that CPs avoid the Case positions reserved for DPs, unless

---

[6] Example (40) is grammatical if the italicized clause follows after a falling intonation, sometimes called a comma-intonation. This would turn it into an instance of right-dislocation.

CPs are part of some concealed DP themselves (see Stowell (1981) and Koster 1978b). Otherwise, CPs are only connected to functional structures as elements specifying DPs.

This little excursion into the distribution of implicit antecedents is perhaps necessary to clarify what kind of free relatives play a role in doubling constructions. Is it the type of pseudocleft sentences (with implicit head) or is it the type of cleft sentences (without implicit head)? As I see no evidence for implicit antecedents in doubling constructions, I will tentatively assume that they involve the free relatives of concealed cleft sentences (but not much hinges on this choice, as pseudocleft sentences will probably also do as the basis for my analysis). Consider therefore once more a cleft sentence like (35a) (repeated here as (41)):

(41)   Het is een boek *dat   hij leest*
            it   is a   book that he reads

It is possible to make a question of this kind of sentence via replacement of the focus (*een boek*) by an interrogative Wh-phrase that is "moved" to the front:

(42)   *Wat*   is het *dat* hij leest?
            what is it that he reads

Interestingly, *dat* in this example can be felicitously replaced by a w-word:

(43)   *Wat* is het *wat* hij leest?
            what is it what he reads

This must be some form of interrogative concord (see Felser 2001), as I have a preference for the d-word in (41) (although judgments tend to vary in that respect):

(44)   ?Het is een boek wat   hij leest
              it   is a   book what he reads

However, a preference for w-words is not limited to questions like (43), it also holds for other constructions supposed to involve A'-movement, like Topicalization:

(45)   Een boek is het wat (?dat) hij leest
            a   book is it what (that) he reads

With a preposed demonstrative *dat*, the w-word is even obligatory:

(46)   Dat is het wat (*dat) hij leest
            that is it   what (that) he reads

As the reader might have guessed, I am particularly concerned with the emergence of proto-doubling as seen in (43). The same pattern can be observed with non-neuter nouns and complex Wh-phrases like *welke man* ('which man'):

(47)   a.   *Wie* is het *wie* jij gezien hebt?
               wie is it who you seen have
               'Who is it that you saw?'

        b.   *Welke man* is het *wie* jij gezien hebt?
               which man is it   who you seen have
               'Which man is it that you saw?'

With one exception, to which I will return, this gives you the full pattern that serves as the basis for the copying constructions under discussion. Note that in all these clefts with questioned focus, only the first Wh-phrase is interrogative. The second Wh-word (the proto-copy) introduces a free relative (possibly without implicit head).

       So, what do we need to derive the doubling constructions under consideration? Very little. The only elements we have to add are phrases like *denk je* ('think you'), *geloof je* ('believe/think you'), etc., which are parentheticals in the simplest cases (Cf. Reis 2000). Other elements that have to go are elements that are easily deleted in Dutch anyway, like forms of the verb *zijn* ('be') and the expletive *het* ('it'). So, on the basis of (43) we can form

(48)   *Wa*t ~~is het~~ denk je *wat* hij leest?
        what ~~is it~~   think you what he reads
        'What do you think that he reads'

Voilà, this gives you all you need to derive questions with doubling! Cleft sentence like those in (47) can also easily be built into the required contexts:

(49)   a.   *Wie* ~~is het~~ denk je *wie* jij gezien hebt?
               who ~~is het~~ think you who you seen have
               'Who do you think that you saw'

        b.   *Welke man* ~~is het~~ denk je   *wie* jij gezien hebt?
               which man ~~is it~~ think you who you seen has
               'Which man do you think that you saw'

This analysis not only makes transparent why, apart from the first interrogative Wh-phrase, the next copies are introducing relative clauses with relative pronouns. It also explains why the alleged copies are not always full copies, thereby in principle explaining an important set of counterexamples to the copying theory. Note that (49b) also is a counterexample to the copying theory of Barbiers *et al.* (2008), which claims that higher chain links can never be more specified than the next links. Example (49b) shows exactly the opposite.

       Now consider the other construction, which does not involve copies but *wat*, which was previously seen as a scope marker by many linguists. Speakers of the relevant idiolect accept the following kind of sentence:

(50)   *Wat* denk je *wie* je gezien hebt?
        what think you who you seen have

The idea that *wat* is a scope marker has always been suspect because *wat* also is a regular Wh-phrase, just like English *what*. Things fall into place if we assume it is just that and also that the sentences in question can be derived in the same simple way as sentences with full

doubling. Thus, what I propose is that, as in the other cases that we have seen, only *what* is an interrogative Wh-phrase, particularly one corresponding to an argument specified by a relative clause, and not some obscure scope marker. This is, in fact the position assumed by the advocates of the Indirect Linking Approach of Felser (2001) and others. My implementation of their idea is that I derive the sentences in question from cleft sentences with Wh-moved focus, as in the examples we have just seen. Thus the structure of (50), I claim, is as follows:

(51)   *Wat* ~~is het~~ denk   je   *wie*   je gezien hebt?
         what ~~is it~~   think you who you seen have

The concealed cleft sentence is slightly odd in standard Dutch, due to a gender mismatch. Thus, while (52a) is perfect, (52b) is not accepted to the same degree. But this is perhaps what one might expect with a phenomenon that is variable across dialects:

(52)   a.   *Wat* is het *wat* je   gedaan hebt?
                what is it what you done have
                'What is it what you did?'

       b.   ?*Wat* is het *wie*   je   gezien hebt?
                 what is het who you seen   have

But note that even in standard Dutch it is sometimes possible to associate a neuter pronoun with a feminine or masculine NP, as in:

(53)   **Dat** is de president
       that is the president

Here we see the neuter demonstrative *dat* ('that'), but we can also have the non-neuter demonstrative *die* ('that') or a full NP *die man* ('that man'):

(54)   a.   **Die** daar is de president
                that there is the president

       b.   **Die man** is de president
                that man is the president

All three cases have a corresponding interrogative form:

(55)   a.   *Wat* is de president?
                 what is the president
                 'What is the president?'

         b.   *Wie* is de president?
                 who is the president
                 'Who is the president?'

         c.   *Welke man* is de president?
                 which man is the president
                 'Which man is the president'

We can come closer to (pseudo)cleft sentences by considering Dutch topicalizations, which typically can involve A'-moved d-words, and which create correspondence problems very similar to what we see with (pseudo)clefts (for the correspondence problem, see Higgins 1973). Particularly with copular predications, as in (52b), we see the neuter form of the d-word *dat* (56a), while when linked to an argument position, we see the non-neuter form *die* (56b):

(56)   a.   De president, *dat* is een schurk
                 the president that is   a   scoundrel
                 'The president is a scoundrel'

         b.   De president, *die* zie je   hier maar zelden
                 the president that see you here only rarely
                 'The president you see only rarely here'

Now note that the DPs at the beginning of the sentence can be replaced by free relatives:

(57)   a.   Wie je   daar ziet, *dat* is een schurk
                 who you there see that is a scoundrel
                 'Who you see there is a scoundrel'

         b.   Wie je   daar ziet, *die* zie je   hier maar zelden
                 who you there see that see you here only rarely
                 'Who you see there, you see here only rarely'

In other words, linking of a free relative clause to a copular predicate via a neuter demonstrative is well-attested, even for free relatives about persons, like (57a).
Applied to the usual range of free relatives that we encounter in questions involving the Indirect Linking Approach, we can think of the following examples, which sound quite reasonable to my ear:

(58) a.   Dat is het wie   hij gezien heeft
            that is it   whom he seen   has
            'That is the one he saw'

     b.   Dat is het   welke jongen hij bedoelde
            that is it   which boy   he meant
            'That is the boy he had in mind'

     c.   Dat is het met wie   hij gesproken heeft
            that is it with whom he talked   has
            'That is the one with whom he talked'

Corresponding to these, we have the following questions:

(59) a.   *Wat* is het *wie*   hij gezien heeft?
            what is it   whom he seen   has
            'What is the one he saw?'

     b.   *Wat* is het *welke jongen* hij bedoelde?
            what is it   which boy   he meant
            'What is the boy he had in mind?'

     c.   *Wat* is het *met wie*   hij gesproken heeft?
            what is it with whom he talked   has
            'What is the one with whom he talked?'

Next, we can build such examples into the contexts that yield questions with apparent doubling of Wh-phrases:

(60) a.   *Wat* ~~is het~~ denk je   *wie*   hij gezien heeft?
            what is it   think you whom he seen   has
            'What do you think he saw?'

     b.   *Wat* ~~is het~~ denk je   *welke jongen* hij bedoelde?
            what is it think you   which boy he   meant
            'Which boy he had in mind?'

     c.   *Wat* ~~is het~~ denk je   *met wie*   hij gesproken heeft?
            what is it think you with whom he talked   has
            'With whom do you think he talked?'

So far, I have assumed that phrases like *denk je* ('think you') are parenthetical in the constructions under discussion. However, that is not necessarily true. Due to normal inversion after Wh-movement in root clauses, the parenthetical *denk je* in (60a) is indistinguishable from instances of the same verb seen not as a parenthetical but an ordinary verb with clausal complement. That there are really two variants can be seen when we embed questions with chains of Wh-phrases under verbs like *afvragen* ('to wonder') (as always, only accepted by speakers of the dialect that allow the construction in question):

(61) a.   Ik vraag me af *wat*, denk   je,   *wie* hij gezien heeft
             I   wonder   what think you who it   seen   has
             'I wonder who, you think, he saw

      b.   Ik vraag me af *wat* je   denkt *wie* hij gezien heeft'
             I   wonder   what you think who he   seen   has
             'I wonder who you think that he saw'

With non-parenthetical *je denkt* ('you think'), the underlying complementizer *dat* ('dat') has
to be deleted as well. This leads to the following structure (with the deleted part as indicated):

(62) Ik vraag me af *wat*   je   denkt ~~dat het is~~ *wie* hij gezien heeft
      I   wonder   what you think ~~that it   is~~ who he   seen   has
      'I wonder who you think that he has seen'

This variant accounts for the well-known occurrence of Wh-questions with apparent multiple
copies of *wat* ('what', was in German):

(63) ***Wat*** denk je   ~~dat het was~~ ***wat*** hij zei ~~dat het was~~ ***wat*** hij geloofde ~~dat het was~~ ***wie***
      what think you ~~that it was~~ what he said ~~that it was~~ what he believed that het was who
      hij gezien had
      he seen   had
      'Who do you think that he said that he believed that he had seen'

Contrary to what is claimed by traditional analyses, such examples do not involve
interrogative scope markers but only one -real- interrogative Wh-phrase, namely the leftmost
wat, which corresponds to an argument in the spirit of the Indirect Dependency Approach.
The other Wh-phrases are not copies but SpecCPs of the specifying free relative clauses of
concealed cleft constructions. Exactly the same analysis applies to examples with what looks
like real copies:

(64) ***Wie*** denk je   ~~dat het was~~ ***wie*** hij zei ~~dat het was~~ ***wie*** hij geloofde ~~dat het was~~ ***wie***
      who think you ~~that it was~~ who he said ~~that it was~~ who he believed that   it was who
      hij gezien had
      he seen   had
      'Who do you think that he said that he believed that he had seen'

Altogether, then, I believe that the alternative sketched so far refutes the copying analysis. In
the next section, I will show that the analysis really solves the many empirical problems of the
extant copying analysis.

4. The problems solved

The empirical problems of the copying theory were discussed above and were summarized in
(17), repeated here for convenience:

(17) a.   Wh-phrases show up in positions where they do not seem licensed
      b.   Wh-phrases are only spelled out in intermediate SpecCPs
      c.   Copying of complex Wh-phrases is prohibited
      d.   Copies are often partial and unequal, with a pattern unexpected
                and unexplained under copying theory

I will show that practically all of these problems disappear under the analysis proposed so far.

4.1 Unlicensed positions

Earlier on, following Schippers (2009) and others, the following data were discussed (18-19, repeated here):

(18) a.   [_{CP1} *Was* meinst du [_{CP2} *wen* Marie geküsst hat?]]
              what think you   who Marie kissed has

      b.   [_{CP1} *Wen* meinst du [_{CP2} *wen* Marie geküsst hat?]]
              who think   you   who Marie kissed has

(19) a.   Er fragt [*wen*   Marie geküsst hat]
              he asks who Marie kissed has

      b.   *Er meint [*wen*   Marie geküsst hat]
              he thinks who Marie kissed has

The problem in these German data (and in similar cases in Dutch) is that a verb like *meinen* ('to think'), does not normally attract an interrogative Wh-phrase, like *fragen* in (19a). So far, there have only been *ad hoc* solutions to this problem, all assuming that the Wh-phrases are actually in the Spec of the CP selected by the verb *meinen*. Under our analysis, however, these facts follow immediately, as the Wh-phrases in question are not interrogative but the Specs of relative clauses. In fact, *wen* in the embedded clauses in (18) is not even in the SpecCP of the CP selected by *meinen* (in the German equivalents of structures (63) and (64):

(65)   *Was* meinst du ~~dass es ist~~ *wen* Marie geküsst hat?
       what think you that it is who Mary kissed   had

Related evidence can be found in the distribution of the interrogative complementizer *of* (*cf.* English 'if') in Dutch. Normally, Dutch indirect questions have an optional *of* following the fronted Wh-phrase:

(66)   Ik vraag me af met wie   (of) hij gesproken heeft
       I   wonder   with whom (if) he talked   has

This optional interrogative complementizer is not expected to occur in free relatives, like those in cleft sentences:

(67)   *Het is Peter [met *wie*   of hij sprak]
        it   is Peter with whom if he talked

This pattern is preserved when the focus of the cleft sentence is questioned:

(68)   ?*Wie* is het [*met wie*   of hij sprak]?
         who is it   [with whom if he talked

Curiously, to my ear at least, (68) is better than (67), but that must be resulting from the kind of interrogative concord that we will also see with Wh-*in situ* and that we saw before (see the discussion of (43) above). Anyway, even for those who tolerate (68) to some degree, the clause between brackets remains a free relative. For those who judge (68) to be unacceptable, the corresponding question (supposed to involve copying) is predicted to be unacceptable as well:

(69)   ?*Wie* denk je *met wie*   *of* hij sprak
         who think you with whom if he talked

Be this as it may, the Wh-phrase *met wie* is not attracted by the verb *denken* ('to think') in examples like (69) under the current analysis.

4.2 Wh-phrases only in intermediate SpecCPs
Since copying plus deletion is supposed to be how Internal Merge works, one expects copies in all positions where traces were thought to be in earlier theories. Unexpectedly, this is far from true. "Visible" copies are only attested for the Specs of CPs, which are only a fraction of the positions formerly believed to be occupied by traces. Consider the following example ((20), repeated here):

(20)   *[CP1 *Was* meinst du [CP2 Marie hat *wen* geküsst?]]
             what think you   Marie has who kissed?

Under standard analyses, *was* is either a scope marker or some Wh-moved argument. In general, an interrogative Wh-phrase in SpecCP is sufficient for the occurrence of Wh-*in situ* in a lower clause, as in the English example (21), repeated here for convenience:

(21)   *Who* said [CP2 that Mary read *what*]?

Similar facts can be observed in German or Dutch, which makes the ungrammaticality of (20) rather mysterious. It is, in fact, unexplained so far. The same mystery pertains to the doubling variant ((22), repeated here):

(22)   **Wen* meinst du [CP2 Marie hat *wen* gesehen]?

Under the alternative theory presented here, facts like (20) and (22) immediately follow: the second copy is in the Spec of a free relative clause and Wh-*in situ* does not exist for relative pronouns:

(70)   *Het is Peter [*met wie*   hij *wie*   gezien heeft]
             it   is Peter with whom he whom seen   has

However, when the focus is questioned, Wh-*in situ* becomes more acceptable, even if the clause remains a relative clause:

(71)   *Wie* is het [met *wie*   hij *wie*   gezien heeft]?
        who is it   with whom he whom seen   has

This is surprising, as it is generally believed that Wh-*in situ* in English, German and Dutch applies to interrogative Wh-phrases only if the first SpecCP up also contains an interrogative Wh-phrase. In (71), the rightmost *wie* is an interrogative Wh-*in situ,* while the clause is headed by *met wie*, which is not interrogative but the Spec of a relative clause. This must be the interrogative concord discussed in the previous section: although *met wie* is introducing a relative clause and therefore is not interrogative itself, it is associated with the leftmost *wie*, which ís interrogative. This association appears to be sufficient to license other Wh-phrases *in situ*. This type of interrogative concord also accounts for the data first discussed by Dayal (1994: 140) ((23), repeated here):

(23)   *Was* glaubst du *wann* Hans *an welcher Universität* studiert hat?
        what think you when Hans at which university   studied has
        'When do you think Hans studied at which university?'

As before, the leftmost *was* is a true interrogative, while the specifying relative clause is introduced by *wann.* Thanks to its being associated with *was*, *wann* can license interrogative phrases like *an welcher Universität.*
        Licensing via interrogative concord is not available for (20) and (22), as these examples do not even have a relative Wh-phrase introducing CP2.

4.3 No copying of complex Wh-phrases
The next fact unexpected under a copying theory is that it does not seem to apply to complex Wh-phrases:

(72)   * *Wessen Buch* glaubst du *wessen Buch* Hans liest?
          whose book think you whose book Hans reads
         Intended: 'Whose book do you think Hans is reading?'

This puzzling fact was recently discussed by Nunes (2004) and van Craenenbroeck (2007) and goes back to McDaniel (1989). Another example from the literature is the one cited by Schippers (2009) ((24) above):

(24)   *[CP1 *Welche Bücher* glaubst du [CP2 *welche Bücher* sie gekauft hat?]]
             which books believe you   which books   she bought has?

What makes this fact even more mysterious is that the corresponding variant with the alleged scope marker *was* is considerably better:

(25)   [CP1 *Was* glaubst du [CP2 *welche Bücher* sie gekauft hat?]]
            what believe you   which   books   she bought has
        'Which books did you believe that she bought?'

Facts like (72) and (24) follow from the fact that free relatives with complex Wh-phrases cannot be used as specifications in cleft sentences:

(73)　*Het zijn *zulke boeken welke boeken* zij gekocht heeft
　　　　it　are such　books　which books　she bought　has

This kind of fact pertains for relative clauses in general: normally, a specifying relative clause is introduced by a Spec that is more "pronominal" than the associated head:

(74)　a.　　*Het boek *welk　boek* zij gekocht heeft
　　　　　　the book which book she bought has

　　　b.　　Het boek *hetwelk*　zij　gekocht heeft [somewhat archaic --JK]
　　　　　　the book it.which she bought　has
　　　　　　'The book which she bought'

　　　c.　　Het boek *dat*　zij gekocht heeft
　　　　　　the book that she bought has
　　　　　　'The book that she bought'

This ungrammaticality of (73) is preserved when the focus constituent is questioned (although not as bad as reported for German and certainly not as bad as (73) and (74a) according to the judgment of both my informants and myself) :

(75)　　??*Welke boeken* zijn het *welke boeken* zij gekocht heeft?
　　　　　which books　are　it　which books　she bought　has

As predicted, this kind of cleft sentence cannot be concealed in the constructions under discussion either:

(76)　　　??*Welke　boeken* denk　je　*welke　boeken* zij gekocht heeft?
　　　　　　which books　think you which books　she bought　has

A further prediction is that variants in which the second copy is reduced to a pronominal relative pronoun are possible. The following cleft sentence is grammatical:

(77)　Het zijn *zulke boeken welke*　zij gekocht heeft
　　　　it　are　such　books　which she bought　has
　　　　'Such books is what she bought'

The corresponding variant with questioned focus is grammatical as well:

(78)　*Welke boeken* zijn het *welke*　zij　gekocht heeft?
　　　　which　books　are　it　which she bought　has

As predicted, this pattern can be built into the doubling constructions under discussion. The following example is grammatical in doubling dialects:

(79)　*Welke boeken* denk　je　*welke*　zij gekocht heeft?
　　　　which books　think you which she bought　has
　　　　'Which books do you think she bought'

In sum, the puzzling prohibition against doubling of complex Wh-phrase only is a mystery under the copying hypothesis. If so-called copies are not copies but introducers of relative clauses associated with a single interrogative Wh-phrase, all facts follow, since they can be reduced to the properties of relative clauses (viz. the requirement of pronominal linking of the relative clause to the specified head or focus constituent).

Given this explanation, the relative grammaticality of complex Wh-phrases in the other construction (introduced by *wat* (Dutch) or *was* (German)) is not immediately expected ((25), repeated here):

(25)   [_{CP1} *Was* glaubst du [_{CP2} *welche Bücher* sie gekauft hat?]]
              what believe you   which   books   she bought has
         'Which books did you believe that she bought?'

The same fact can be observed in the relevant dialects of Dutch:

(80)   *Wat*   geloof je *welke boeken zij* gekocht heeft?
         what believe you which books   she bought has
         'Which books did you believe that she bought?'

Here, surprisingly, we do find a complex Wh-phrase introducing the specifying relative clause. However, as we concluded earlier on, *wat* is not directly associated with the Spec of the relative clause, but with the entire clause (which can correspond to the neuter form *het* (*dat*) and the question variant *wat*). Thus the following are not all that bad in doubling dialects (particularly not with comma intonation as indicated by the comma, which turns the free relatives into right-dislocations):

(81)   a.   *Dat* is (het), [*welke boeken zij gekocht heeft*]
                   that is   which books   she bought has

         b.   *Wat* is (het), [*welke boeken zij gekocht heeft*]?
                   that is   which books   she bought has

Note also that we can also refer to a left-dislocated free relative clause of this type with the neuter form of the d-word *dat*:

(82)   *Welke boeken* zij gekocht heeft, *dat* is een geheim
         which books   she bught has   that is a   secret

Like what we see with all clauses, we can also specify the free relative in question with another -appositive- relative, introduced by *what* (compare English *which*):

(83)   Wij zagen *welke boeken zij gekocht had*, *wat* een mooie collectie was
         we   saw   which books she bought had which a nice collection was
         'We saw which books she had bought, which was a nice collection'

In short, it is a well-established fact that free relatives introduced by complex Wh-phrases can be referred to by neuter pronouns like *dat* ('dat'') and *wat* ('wat'). This leads to questions like (81b), which can be built into doubling constuctions, deriving (80) (repeated here), as desired:

(84)   *Wat*   geloof je *welke boeken zij* gekocht heeft?
         what believe you which books   she bought has
         'Which books did you believe that she bought?'

Summarizing, then, I believe that a theory built on specifying relative clauses explains the mysteries of complex Wh-phrases, which remained unresolved under a copying theory. I will now turn to the intriguing distribution of non-equal copies which, I hope to show, form the strongest evidence against copying theories, including theories assuming partial copying.

4.4 Partial and unequal copies
The biggest problem for copying theories is that, more often than not, copies are not real copies but Wh-phrases of considerably different shape. This is not expected under a copying theory since, normally speaking, a copy tends to be an exact replica of the thing copied. This is particularly true if copying is believed to originate from Internal Merge. Solutions found in the literature are usually *ad hoc*, like the notion of "partial copying" proposed by Barbiers *et al.* (2008). Their solution was shown to give undesirable results, as it involved serious island violations and just the opposite result for chain links from the standard one ((12) above instead of the Anaphoric Principle (6)). I would now like to show that the data discussed by Barbiers *et al.* exactly fit the hypothesis that I propose. To see this, let us return for a while to their facts ((9) - (11), repeated here):

(9)   *Wat* denk je   *wie* ik gezien heb?
         what think you who I   seen have

(10)   *Wie* denk je *die*   ik gezien heb?
         who think you rel.pron I   seen have

(11)   *Wat* denk je *die*   ik gezien heb?
         what think you rel.pron I seen have

According the theory proposed here, only the first Wh-phrase is an interrogative (as is also proposed by Den Dikken and Bennis 2009). The second SpecCP is a relative pronoun, as Barbiers *et al.* and Den Dikken and Bennis assume as well. However, all maintain the notion of copying, which is not corroborated by these data but undermined. Like Barbiers *et al.*, Den Dikken and Bennis propose a form of partial copying, claiming that having *die* as a copy (rather than *wie*) is more economical. Such speculations are neither necessary nor sufficient to account for the data, as we will see as soon as we consider a broader class of "copies."
        First of all, note that the facts ((9) – (11)) immediately follow from the theory presented so far. Since this theory claims that second and subsequent copies introduce relative clauses, variants with the relative pronoun *die* are expected, as in the corresponding cleft sentences with questioned focus. Thus, both the following are possible:

(85)   a.   *Wie* is het *wie* ik gezien heb?
                   who is it   who I   seen have

         b.   *Wie* is het *die*   ik gezien heb?
                  who is it   who I   seen have

The second variant is built into (10), The other variant, with *wat*, follows straightforwardly from our general account of *wat*-sentences. This way we can avoid the notion partial copying and its *ad hoc* violations of general principles. In a minute, I will give more examples showing that IM with partial copying is impossible, but let us first have a look at the examples that must be excluded according to Barbiers *et al* .((13) – 15), repeated here):

(13)   *__Wie__ denk   je   *wat* ik gezien heb?
          who think you what I   seen have

(14)   *__Die__   denk   je   *wie* ik gezien heb?
          rel.pron think you who I   seen have

(15)   *__Die__   denk   je   *wat* ik gezien heb?
          rel.pron think you what I seen have

The sentence (14) and (15) are immediately excluded because the leftmost phrase of series of "copies" must be an interrogative Wh-phrase. The first example, (13), is also excluded because non-neuter focus constituents can never be specified with a relative clause introduced by *wat*:

(86)   a.   Het is *die*   daar *wie* ik gezien heb
                 it   is that there who I   seen have
                 'It is that one (person) that I saw'

        b.   *Het is *die*   daar   *wat* ik gezien heb
                 it   is that there what I   seen have

This pattern is preserved when the focus is questioned, as in (87), corresponding to (86b):

(87)   *__Wie__ is *wat*   ik gezien heb?
                   who is what I   seen   have

Failing agreement causes ungrammaticality here and it is not remedied when the structure is built into the construction under discussion, as shown by (13). In short, the data of Barbiers *et al.*, follow in a natural way, without *ad hoc* extensions of the theory.

The case against partial copying can even be further strengthened. Consider the examples (33) above (and repeated here):

(33)   a.   ?*Met wie*   zei   je   *met wie*   je gesproken had?
                 with whom said you with whom you talked   had
                 'With whom did you say that you talked?'

        b.   *Wie*   zei je   *met wie*   je gesproken had?
                 who(m) said you with whom you talked   had

As we observed, every speaker who accepts doubling at all finds (33b) better than (33a). Partial copying (IM of *wie* applied to the subphrase *wie* of the PP) would violate the very strong islandhood of PPs in such cases in Dutch. But even if this impossible extraction from PPs would be allowed for once, it would be unexpected that copying of a subphrase leads to a

better result than full copying, since full copying seems to be the default form of copying. On the assumption that the second copy introduces a relative clause, the facts fall into place. First, consider cleft sentences corresponding to (33a) and (33b):

(88) a.   ?Het is **met die man** *met wie*   hij gesproken heeft
         it   is with that man with whom he talked   has

    b.   Het is **die man** *met wie*   hij gesproken heeft
         it   is that man with whom he talked   has

Even at this level, it is better to have a DP antecedent for the relative clause (88b) than one included in a PP (88a). The difference is preserved when the focus constituents are questioned:

(89) a.   ?*Met wie*   is het *met wie*   hij gesproken heeft
         with whom is it with whom he talked   has

    b.   *Wie*   is het *met wie*   hij gesproken heeft
         whom is it with whom he talked   has

The difference is preserved once more when these sentences are built into the constructions under discussion. So, while this pattern of data is incompatible with all copying theories I know of, it is completely regular in a theory based on specification with relative clauses.
    At least as interesting is the case of "partial copying" briefly discussed as (31):

(31)   *Hóeveel*   zeg je *hoeveel varkens* je   gezien hebt?
     how.many say you how.many pigs   you.PL seen have
     'How many pigs do you say (that) you saw?'

As discussed, this is a strong counterexample to partial copying via IM, as Dutch has a strong Left Branch Condition for such cases. However, the analysis I have in mind can easily accommodate this case. A corresponding cleft sentence would be as follows (since this is about quantity and not about countable units, I prefer *much* over *many* as the correct translation):

(90)   *Zóveel*   is het *hoeveel   varkens* je gezien hebt!
     so.much is it how.much pigs   you seen have

Questioning the focus (*zóveel*) produces:

(91)   *Hóeveel*   is het *hoeveel   varkens* je gezien hebt?
     how.much is it   how.much pigs you seen have

Built into our construction, this yields (31). Again, we see a straightforward pattern, without any island violations, where a copying theory meets insurmountable problems. However, things are even worse for the copying theory if we realize that the opposite order for the associated elements of (31) is also possible (and even slightly better in my judgment):

(92) *Hóeveel varkens* denk je *hoeveel* je gezien hebt?
      how.many pigs   think you how.many you seen have
      'How many pigs did you think (that) you saw?'

Assuming that IM applies from the bottom up, there is no reasonable way to turn the lower *hoeveel* into the larger phrase *hoeveel varkens* of the higher clause, as the pigs would have to be created *ex nihilo*. I am very curious indeed how the proponents of the copying theory would fulfill this miraculous task. Under the analysis proposed here, however, (92) is unproblematic. The corresponding cleft is (93a) and the same with questioned focus is (93b);

(93) a. *Zóveel varkens* zijn (het) *hoeveel* je gezien hebt
          so-much pigs   are it how.much you seen have

      b. *Hóeveel varkens* zijn (het) *hoeveel*   je gezien hebt
          how-much pigs   are   it how.much you seen have

Another example I will discuss was suggested to me by Jeroen van Craenenbroeck (p.c.). There seems to be the following contrast (although according to the judgment of both my informants and myself not as strong as claimed):

(94) a. *Welke jongen* denk je *wie* ik gezien heb
          which boy   think you who I seen have
          'Which boy do you think I saw'

      b. ?*Wie* denk je *welke jongen* ik gezien heb?
           who think you which boy   I seen have

Once more, it is unlikely that we see copies here, because the two associates are so different. The difference in grammaticality follows from what we discussed earlier with respect to complex Wh-phrases: a specifying relative clause must be introduced by a Wh-phrase that is sufficiently pronominal. The underlying clefts with questioned focus phrase corresponding to (94a) and (94b) are as follows:

(95) a. *Welke jongen* is het *wie* ik gezien heb?
          which boy   is it who I seen have

      b. ?*Wie* is het *welke jongen* ik gezien heb
          who is it which boy   I seen have

Building the clefts into our construction, leads to (94), all again as expected. For a copying analysis, however, I see no solution for (94), neither as an account for the grammatical (94a) nor as a basis for the explanation of the less grammatical (94b). Furthermore, (94a) raises the same problems as the previous example about pigs: it is impossible to copy *wie* in (95a) via IM without adding lexical material (both *welke* and *jongen*) and getting rid of *wie* somehow. This is too much, even for a sign of perfection like IM.

My last example would involve a complete metamorphosis under a copying analysis. The case of doubling I have in mind is entirely acceptable in my judgment:

(96) *Welke auteur* denk je *wiens boeken* je wilde   lezen?
      which author think you whose books you wanted read
      'Which author's books did you think you wanted to read?'

No form of copying will work here, neither total nor partial, as the two Wh-phrases are completely different. The relative acceptability of (96) also shows that the ban on the doubling of complex Wh-phrases (as discussed by McDaniel, Nunes and van Craenenbroeck) is too strong. Under the approach advocated here, hower, this example is predicted to be grammatical. Consider again the underlying cleft sentence:

(97) Het is *díe auteur wiens boeken* je wilde   lezen!
      it   is thát author whose books you wanted read
      'It is thát author whose books you wanted to read!'

Questioning the focus leads to the following result:

(98) *Welke auteur* is het *wiens boeken* je wilde   lezen?
      which author is it   whose books you wanted read

Building it into the familiar structures works as before, yielding the acceptable (96):

(99) *Welke auteur* ~~is het~~ denk je *wiens boeken* je   wilde lezen
      which author is it think you whose books you wanted read

Copying would have to go beyond supplementary creation *ex nihilo* to bring about this rather substantial metamorphosis. That seems asking to much.

       All in all, then, we must conclude that as soon as a broader class of data is considered than the usual, copying theory simply does not work for the doubling of Wh-phrases in German and for the doubling dialects of Dutch (including Afrikaans, I suspect). Neither copies nor scope markers (with the form of a Wh-phrase) are found in these languages. In contrast, a non-copying theory based on specification by relative clauses accounts for most of the facts and predicts several interesting new ones.

5. Conclusion

Syntax defines pathways for information exchange. Examples of information exchange are completions and additions. The properties of displacement (formerly "movement") fall into the pattern of completions, i.e., relations with the properties of the Configurational Matrix ((4) above) and unified under the operation "share property." Information exchange is driven by a *difference* in feature content between the interacting categories. For completions, this is expressed by the Anaphoric Principle (6).

       In fact, the Anaphoric Principle is an instantiation of an even deeper principle that also applies to the other major family of syntactic exchanges, namely additions. Views on what additions are differ, but they minimally encompass predications and specifications (see Den Dikken 2006 for discussion). I will not go into the semantic content of predication and specification but I will limit myself to a syntactic similarity and a difference between these two forms of addition. The similarity is in the direction of the information flow, which is just

the opposite from what we see in completions (Principle of Natural Word Order (1); see (5) above):

(100)   **Principle of Natural Word Order (2)**
        Predications and specifications follow the category they add information to

Thus, with completions the "richer" category is on the left, while with specifications it is on the right. A predication (here seen as an "aboutness" relation) usually involves a DP and an XP in the order DP-XP, in accordance with (100). Predication follows the strictest form of locality available, namely sisterhood (DP and XP must be sisters).

Specifications do not necessarily involve a DP as first term, but they often do. Locality for specifications is less strict than for predications and for completions, but does meet the Right Roof Constraint (see Koster 2000 for discussion). It is not my intention here to give a full account of major syntactic interaction types, but I briefly introduce these other types here to show that the Anaphoric Principle does not stand alone but is an instantiation of an even deeper principle of language, namely the following:

(101)   **Syntactic Inertness Principle**
        Syntactic categories do not interact unless they differ

Thus, take a typical specification, like the equatives discussed by Ross (1969) (*a golden igloo* specifies *something beautiful*):

(102)   They saw *something beautiful*: *a golden igloo*

The specification relation works as such because *a golden igloo* differs from *something beautiful* and adds something to it. Suppose now that we replace *a golden igloo* by a copy of *something beautiful*:

(103)   They saw *something beautiful*: *something beautiful*

Maybe this sentence can be interpreted with some effort, but one can be sure that in that case a subtle difference will be imposed on the two instances of *something beautiful*. As also entailed by the Anaphoric Principle, exact copies do not lead to interpretable interactions at all and are therefore, if you will, rejected and sent to hell at the conceptual-intentional interface. Principle (101) is an economy principle, somewhat reminiscent of a Gricean relevance maxim. Given the obvious truth of (101) and in the light of the empirical results of this article, we can be pretty confident that (104) is true as well:

(104)   No syntactic category ever has a copy!

Jan Koster
J.Koster@rug.nl   10 Nov 2009

Bibliography

Barbiers, Sjef, Olaf Koeneman & Marika Lekakou. 2008. Syntactic doubling and the structure of chains. In Charles Chang & Hannah Haynie, eds., *Proceedings of the 26th West Coast conference on formal linguistics. Cascadilla Proceedings Project*. Somerville, Mass.: 77-86.

Blom, Alied & Saskia Daalder. 1977. *Syntaktische theorie & taalbeschrijving*. Muiderberg: Dick Coutinho.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky, eds., *A Festschrift for Morris Halle*. New York, etc.: Holt, Rinehart & Winston: 232-286.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser, eds., *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.: 1-52.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner, eds., *Interfaces + Recursion = Language?* Berlin: Mouton de Gruyter: 1-29.

Cinque, Guglielmo. 1983. Island effects, subjacency, ECP/connectedness and reconstruction. Ms.Venice: Università di Venezia.

Cinque. Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.

Craenenbroeck, Jeroen van. 2007. Complex wh-phrases don't move. On the interaction between the split CP hypothesis and the syntax of wh-movement [to appear in *CLS* 43].

Dayal, Veneeta. 1994. Scope marking as indirect wh-dependency. *Natural Language Semantics* 2: 137–170.

Dikken, Marcel den. 2006. *Relators and linkers: the syntax of predication, predicate inversion and copulas*. Cambridge, Mass.: MIT Press.

Dikken, Marcel den. 2009. On the nature and distribution of successive cyclicity: adjunction, resumption, and scope marking as the roads to success in long-distance relation building. Ms. New York: Linguistics Program, CUNY Graduate Center.

Dikken, Marcel den & Hans Bennis. 2009. Wat dacht je wat? On the strategies for the formation of long-distance dependencies in constituent questions. Amsterdam: Meertens Institute. [KNAW Visiting Professors Lecture Series, 15 October 2009].

Du Plessis, Hans. 1977. Wh-movement in Afrikaans. *Linguistic Inquiry* 8, 723-726.

Fanselow, Gisbert & Damir Ćavar. 2001. Remarks on the economy of pronunciation. In Gereon Müller & Wolfgang Sternefeld, eds., *Competition in syntax*. Berlin: Mouton de Gruyter: 107-150.

Felser, Claudia. 2001. Wh-expletives and secondary predication: German partial wh-movement reconsidered. *Journal of Germanic Linguistics* 13: 5-38.

Felser, Claudia. 2004. Wh-copying, phases, and successive cyclicity. *Lingua* 114: 543-574.

Heycock, Caroline & Anthony Kroch. 1999. Pseudocleft connectedness: implications for the LF interface level. *Linguistic Inquiry* 30: 365–397.

Hiemstra, Inge. 1986. Some aspects of wh-questions in Frisian. *Nowele* 8: 97–110.

Higgins, Roger. 1973. The pseudo-cleft construction in English. Doctoral dissertation, MIT, Cambridge, Mass. [Published by Garland, New York, 1979]

Hornstein, Norbert, Jairo Nunes & Kleanthes Grohmann. 2005. *Understanding minimalism*. Cambridge: Cambridge University Press.

Kallulli, Dalina. 2009. Why and how to distinguish between pro and trace. Ms., Vienna: University of Vienna.

Kayne, Richard. 1994. The antisymmetry of syntax. Cambridge, Mass.: MIT Press.

Koster, Jan. 1978. Why subject sentences don't exist. In Samuel J. Keyser, ed., *Recent transformational studies in European languages.* Cambridge, Mass.: MIT Press.

Koster, Jan. 1987. *Domains and dynasties: the radical autonomy of syntax*. Dordrecht: Foris.

Koster, Jan. 2000. Extraposition as parallel construal. Ms., Groningen: University of Groningen [see: http://www.let.rug.nl/koster/1999.htm].

Koster, Jan. 2003. All languages are tense second. In Jan Koster & Henk van Riemsdijk, eds., *Germania et alia: a linguistic webschrift for Hans den Besten*. Groningen: University of Groningen [http://www.let.rug.nl/~koster/DenBesten/contents.htm].

Koster, Jan. 2008. A reappraisal of classical V2 and scrambling in Dutch. In *Groninger Arbeiten zur germanistischen Linguistik* 46: 27-54.

Lutz,Uli, Gereon Müller & Arnim von Stechow. 2000, eds., *Wh-scope marking*. Amsterdam and Philadelphia: John Benjamins.

McDaniel, Dana. 1989. Partial and multiple Wh-movement. *Natural Language and Linguistic Theory* 7: 565-604.

Moro, Andrea. 1990. There-raising: principles across levels. Paper presented at the 13th GLOW Colloquium, Cambridge, England.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, Mass.: MIT Press.

Obenauer, Hans-Georg. 1984. On the identification of empty categories. *The Linguistic Review 4*: 153-202.

Ouhalla, Jamal. 2001. Parasitic gaps and resumptive pronouns. In Peter Culicover & Paul Postal, *Parasitic gaps.* Cambridge, Mass.: MIT Press.: 147-179.

Postal, Paul. 1993. Parasitic gaps and the across-the -board phenomenon. *Linguistic Inquiry* 24: 735-754.

Riemsdijk, Henk van. 1978. *A case study in syntactic markedness: the binding nature of prepositional phrases.* Foris: Dordrecht.

Riemsdijk, Henk van. 1983. Correspondence effects and the Empty Category Principle. In Yukio Otsu, Henk van Riemsdijk, K. Inoue, A. Kamio & N. Kawasaki, eds., *Studies in generative grammar and language acquisition: a report on recent trends in linguistics.* Tokyo: International Christian University: 5-16.

Reis, Marga. 2000. On the parenthetical features of German was ... w constructions and how to account for them. In Lutz, Müller, & von Stechow, eds.: 359–407.

Ross, John Robert. 1969. Adjectives as noun phrases. In David Reibel & Sanford Schane, eds., *Modern studies in English: readings in transformational grammar*. Englewood Cliffs, NJ.: Prentice-Hall.

Schippers, Ankelien. 2009. Some people are repeaters: medial copy spell-out in long-distance wh-dependencies. Ms., Groningen: University of Groningen.

Stepanov, Arthur & Penka Stateva. 2006. Successive cyclicity as residual wh-scope marking. *Lingua* 116: 2107-2153.

Stowell, Tim. 1981. Origins of phrase structure. Doctoral dissertation, MIT, Cambridge, Mass.

Vergnaud, Jean-Roger. 1974. French relative clauses. Doctoral dissertation, MIT, Cambridge, Mass.

Wasow, Thomas. 1972. Anaphoric relations in English. Doctoral dissertation, MIT, Cambridge, Mass.

Williams, Edwin. 1994. *Thematic structure in syntax*. Cambridge, Mass.: MIT.