

Copy what?

Annelien Schippers

1. Introduction

The existence of wh-copy constructions, here illustrated for Dutch in (1) below, has been taken to support the hypothesis that long-distance movement proceeds successive-cyclically, and leaves behind copies during this process.

- (1) [CP₁ Wie denk je [CP₂ wie Jan de beste taalkundige vindt?]]
who think you who Jan the best linguist finds
'Who do you think Jan considers to be the best linguist?'

However, the assumption that the intermediate wh-phrase in (1) is a true copy of the higher wh-phrase raises many problematic issues (cf. Schippers, 2009; Koster, 2009). Moreover, there are cases in which the intermediate wh-phrase cannot in all honesty be considered a copy of the higher wh-phrase.

In this paper, I present a survey I carried out concerning natural occurrences of the wh-copy construction in Dutch, using the search engine Google. As it turns out, some of these examples cannot be explained under the copy theory of movement. In fact, they cannot be accounted for under any theory of (long-distance) extraction. This thus provides evidence for the position taken by Koster (2009), which holds that the wh-phrases in wh-copy constructions cannot be true copies of each other, but must instead be regarded as independent lexical items.

2. The data

I conducted a Google search focusing on wh-copy constructions where complex wh-phrases consisting of a wh-part and a postposition (R-pronoun) are long-distance moved. This type of wh-phrase is of interest for two reasons. First of all, it has often been noted that more complex wh-phrases do not allow copying (cf. Fanselow & Mahajan, 2000; Höhle, 2000). It has been proposed that this is due to the fact that the wh-phrase undergoes fusion with (or cliticizes onto) C, and that this is only possible when the wh-phrase is pronominal, and not when it is phrasal (cf. Fanselow & Mahajan, 2000; Nunes, 2004).

A second reason why postpositional wh-phrases are of interest is because they allow p-stranding (cf. Van Riemsdijk, 1978). Because of this, several copying variants are predicted to be possible: either the entire complex consisting of the wh-phrase plus postposition may get copied, or the postposition may be stranded in base position, resulting in copying of only the wh-phrase itself. Another possibility that presents itself is stranding of the postposition in an intermediate SpecCP, which has been reported to occur in Afrikaans and illustrated here in (2) below (example from Du Plessis, 1977; the base-copy is in strikethrough).

(2) **Waar/wat** dink julle **voor** werk ons ~~waarvoor~~?

Where/what think you for work us

'Where do you think we work for?'

The Internet search discussed below indeed returned examples that suggest all such variants are possible in Dutch. Before discussing these data in section 2.2, I shortly discuss the procedure by which I collected these examples in section 2.1 below.

2.1 Method

Dutch corpus studies on long-distance wh-movement constructions show that they predominantly show up with the matrix verb *denken* ‘think’ and a second person matrix subject (cf. Verhagen, 2006; Schippers & Hoeksema, 2009). The search term for Google I used therefore consisted of the present tense finite second person form of the matrix verb *denken* ‘think’. For the matrix subject position, I used a wildcard. In (3) a, b, c and d, the search terms that were used are illustrated:

- (3) a "Wh denk * Wh * PP" = stranding in base position, copying of wh-phrase only
b "WhPP denk * WhPP" = full copying of Wh-phrase and postposition
c "Wh denk * WhPP" = Wh in matrix SpecCP, Wh+PP in subordinate SpecCP
d "WhPP denk * Wh" = Wh+PP in matrix SpecCP, Wh in subordinate SpecCP

Note that I did not search specifically for cases of p-stranding in the intermediate SpecCP. The reason for this is mainly a practical one, since such a search term returned a huge number of hits involving root questions with preposition stranding, and it would be too cumbersome to inspect all the relevant data non-automatically. Moreover, as I will discuss later, there is reason to think that there is no such thing as preposition stranding in intermediate SpecCP, not even in Afrikaans.

2.2 Results

In total, the search terms listed in (3) returned 107 examples. It turned out that the majority of these examples concerned preposition stranding in base position, with copies of the wh-

phrase in both SpecCPs (77 occurrences, 72% of total). One of the attested examples is illustrated in (4).

- (4) [CP1 Waar denk je [CP2 waar de wegenbelasting heengaat?]]
where think you where the road.tax to.goes
'Where do you think the road tax goes to'

The next largest group was formed by cases where the entire complex Wh+PP was copied (20 occurrences, 19% of total). Such cases were almost exclusively restricted to the phrases *waarom* 'why' and *wanneer* 'when'. These phrases behave like monomorphemic words and generally do not allow stranding (especially true for *wanneer*), explaining why the full adverb is copied in these cases. There was in fact only one case in which an adverbial wh-phrase that readily allows stranding was doubled, shown in (5):

- (5) [CP1 Waarvan denk je [CP2 waarvan ik hou?]]
Where.of think you where.of I like?
'What do you think I like?'

So far, the hypothesis that complex wh-phrases cannot be copied seems to be confirmed by the data. However, various examples where complex wh-phrases are copied have been cited in the literature. In (6) is a German example taken from Pankau (2009) and in (7) an Afrikaans example from Du Plessis (1977).

- (6) [CP1 Mit wem glaubst du [CP2 mit wem sie sich getroffen hat?]]
With whom think you with whom she self met has

‘Whom do you think she has met?’

(7) [CP₁ Met wie het jy gesê [CP₂ met wie gaan Jan trou?]]

With whom have you said with whom go John marry

‘With whom did you say John will marry?’

There is thus no reason to assume that copying of complex wh-phrases itself is banned, and consequently there is also no evidence for the hypothesis that the intermediate copy undergoes fusion with (or cliticizes onto) C. It appears that the only type of complex wh-phrases that really cannot be copied concern wh-phrases with an NP restrictor (i.e. ‘which + NP’ phrases). In effect, while it may very well be true that complex wh-phrases are more marked than pronominal wh-phrases when it comes to wh-copying, they are certainly not excluded from showing up in this construction. More specifically, it appears that the right distinction between wh-phrases that may and may not show up in intermediate SpecCPs in the wh-copy construction is determined by whether the wh-phrase in question can be an introducer of a free relative clause (cf. Koster, 2009; Pankau, 2009). This explains why prepositional wh-phrases like *mit wem* ‘with whom’ are allowed, but ‘which + NP’ phrases are not, since the latter cannot introduce a free relative, while the former can.

The examples discussed so far do not provide evidence against a copy analysis of the constructions under consideration. However, the constructions that are of particular concern are those of the type in (3) c and d, which would involve unequal ‘copies’. Interestingly, I have indeed been able to find such examples, although they appear much less frequent than the identical copy cases discussed earlier. Among the examples I found using the search term in (3) c were ones like in (8) below.

(8) [CP1 Waar denk je [CP2 waarmee je bezig bent?]]

Where think you where.with you busy are

'What do you think you're doing?'

In total, I found 6 of these examples (5.6% of total). This type of construction looks reminiscent of the Afrikaans example in (2), where the postposition is stranded in an intermediate SpecCP. However, it is unlikely that constructions as in (8) can be explained along the lines of a stranding in SpecCP analysis. For this to go through, it would have to be assumed that not only the postposition, but also the wh-part is stranded. This is something which otherwise does not seem to be allowed in Dutch, i.e. stranding always leaves behind only the postposition, never the wh-part (as well). Even more problematic is the fact that the construction type in (8) is attested with the wh-phrase *waarom* 'why', here illustrated in (9) below:

(9) [CP1 Waar denk je [CP2 waarom ik drank heb leren drinken?]]

where think you why I alcohol have learned drink

'Why do you think I learned to drink alcohol?'

As noted before, *waarom* 'why' seems to behave like a monomorphemic word, and usually does not strand the postposition. This example therefore casts further doubt on the assumption that examples as in (8) and (9) involve some kind of stranding in SpecCP. This in turn suggests that (8) and (9) do not involve (partial) copying. Moreover, as pointed out in Rackowski & Richards (2005) and Den Dikken (2009), there is reason to believe that the Afrikaans example in (2) does not involve p-stranding in the intermediate CP either, since constructions as in (2) are excluded once an extra embedding is added.

Next, I would like to turn to the constructions I found using the search term in (3) d, which would form the strongest evidence against a copy analysis. Although such examples were again very infrequent, I was in fact able to find 4 of them (3.7% of total). One of the relevant examples is shown in (11):

- (11) [CP₁ Waarvoor denk jij [CP₂ waar deze mensen voor dienen en voor worden betaald?]]
where.for think you where these people for serve and for to be paid
'For what do you think these people serve and are being paid'

Again, these examples cannot be explained by assuming the wh-phrases in SpecCP are copies of each other: since the wh-phrase in the embedded SpecCP has already stranded the postposition, there is no source for it in the highest SpecCP. If a copy analysis is maintained, one would have to jump some hoops in order to explain examples like (11). Even if one is willing to do so, there is one final piece of evidence that I believe argues decisively against a copy analysis. Namely, next to cases like (11), I also found two examples where the postposition *waarvoor* 'what for' was stranded in the lower SpecCP, while the highest SpecCP is occupied by an altogether different wh-phrase: *waarom* 'why'. One of these cases is given in (12) below:

- (12) [CP₁ Waarom denk je anders [CP₂ waar die voor bedoeld zijn?]]
why think you otherwise where those for meant are
'Where do you otherwise think those are meant for?'

I do not think that anyone would want to conjecture in all fairness that examples like (12) involve copying. The existence of constructions as in (12) therefore causes serious problems for a copy analysis.

3. No more copying?

The data discussed in the previous section provides evidence against a copy analysis and in favor of an analysis in which the higher and lower wh-phrase are considered to be two distinct items. Such analyses have recently been proposed in Den Dikken (2009) and Koster (2009). Den Dikken suggests copy constructions are a variant of so-called partial wh-movement constructions, exemplified in (13).

(13) [_{CP1} Wat denk je [_{CP2} wie Marie gekust heeft?]]

What think you who Marie kissed has

'Who do you think Marie kissed?'

He argues that both wh-copy and partial wh-movement constructions have a scope marker (*wat* 'what') that is base generated in the matrix clause. In case of partial wh-movement, this scope marker enters into an indirect dependency with the wh-phrase in the embedded CP, while in the wh-copy constructions the scope marker enters into a direct dependency with the embedded wh-phrase. This direct dependency facilitates concord between the scope marker and the lower wh-phrase, whereby the scope marker may copy over features of the lower wh-phrase. If it copies all the features (full concord), only the highest wh-phrase is spelled out, since in this case the lower and higher wh-phrase are non-distinct. The resulting output then

resembles a standard long-distance extraction construction. Next to full concord, Den Dikken argues it is also possible to have partial concord. In that case, the scope marker only copies part of the features of higher wh-phrase (e.g. its D-features). This results in surface-identical wh-phrases, but crucially, their featural make-up is different, which means the wh-phrases in question are also distinct. Therefore, both the lower and higher wh-phrase are spelled out, and the derivation surfaces as a wh-copy construction. Den Dikken's analysis departs from standard analyses of wh-copying in that the lower and higher wh-phrase are not considered to be true (movement) copies of each other. On the other hand, his analysis still has a strong 'copying' flavor to it, due to the fact that concord between the scope marker and wh-phrase obtains. One of the consequences of this is that the higher wh-phrase may never have features lacking from the lower wh-phrase. For this reason, his analysis runs into trouble when it comes to examples like (11) and (12), where the highest wh-phrase clearly has features lacking from the lower one.

This problem does not arise for the analysis proposed in Koster (2009). He also argues that partial wh-movement and wh-copy constructions are similar in that both involve base generation of the higher wh-phrase in the matrix clause itself. But rather than assuming wh-copy constructions are a type of scope marking constructions, he assumes the matrix clause is a kind of cleft sentence, and the embedded clause essentially a type of free relative. The two clauses are linked to each other because the focus of the cleft sentence questions the embedded wh-phrase (which in Koster's analysis is a free relative pronoun). The underlying representation of a copy construction in his analysis looks like something as in (14) below, repeated from (1):

(14) [Wie (is het) denk je [wie Jan de beste taalkundige vindt?]]

who is it think you who Jan the best linguist finds

Koster assumes the copula *is* 'are' and expletive *het* 'it' are deleted, and that the matrix verb and subject are parenthetical. Only the highest wh-phrase is considered to be an interrogative: the embedded wh-phrase is a type of relative pronoun, which happens to be spelled out as a wh-word due to interrogative concord. Hence, just as Den Dikken, Koster assumes some form of concord is at work, but contrary to Den Dikken, this concord is restricted to interrogativity only. Koster's analysis is thus able to account for the fact that in some cases, the wh-phrases in SpecCP are not (partial) copies of each other. Hence, contrary to Den Dikken, he is able to account for constructions as in (11) and (12) above.

One may object that such cases where 'copies' are unequal are quite infrequent, and should hence be disregarded as errors. I do not think this is the case. First of all, examples of 'imperfect copying' have been attested in the literature before (cf. Fanselow & Ćavar, 2001, Anyadi & Tamrazian, 1993, Pankau, 2009), and they are also particularly frequent in child language data (cf. Thornton, 1990; Van Kampen, 1997). And although the data discussed here are not the most reliable ones, they do suggest that such 'imperfect copying' cases exist in Dutch as well. The most conservative conclusion that thus can be drawn from these data is that not all cases of wh-copying are the result of successive-cyclic movement through SpecCP, and that there is indeed a legitimate need for an analysis along the lines of Koster (2009). The subsequent question is whether all cases of wh-copying should be subsumed under such an analysis. I will leave this issue open for further discussion.

In conclusion, the phenomenon of imperfect copying appears to be real, and is not compatible with a long-distance extraction analysis of the constructions under consideration. What remains to be decided is whether an analysis along the lines of Koster (2009) must be extended to all cases of wh-copying.

References

- Anyadi, Stefanie and Armine Tamrazian (1993). Wh-movement in Armenian and Ruhr German. *UCL Working Papers in Linguistics* 5, 1 - 22.
- Den Dikken, Marcel. (2009). On the distribution of successive-cyclicity – Adjunction, resumption and scope marking as the roads to success in long-distance relation building. Ms., New York.
- Du Plessis, Hans. (1977). Wh-movement in Afrikaans. *Linguistic Inquiry* 8, 723 - 726.
- Fanselow, Gisbert and Damir Ćavar (2001). Remarks on the economy of pronunciation. In Gereon Müller & Wolfgang Sternefeld (eds.), *Competition in syntax*, pp. 107 - 150. Berlin: Mouton de Gruyter.
- Fanselow, Gisbert & Anoop Mahajan (2000). Towards a Minimalist Theory of Wh-Expletives, Wh-Copying, and Successive Cyclicity. In Uli Lutz, Gereon Müller and Arnim von Stechow (eds.), *Wh-Scope Marking*, pp. 195 - 230. Amsterdam: John Benjamins.
- Höhle, Tilman N. (2000). The W-... W- Construction: Appositive or Scope Indicating. In Uli Lutz, Gereon Müller and Arnim von Stechow (eds.), *Wh-Scope Marking*, pp. 249 - 270. Amsterdam: John Benjamins.
- Kampen, Jacqueline van (1997). *First steps in wh-movement*. Ph.D. dissertation, Utrecht University.
- Koster, Jan (2009). IM not perfect. Ms., University of Groningen, Groningen.
- Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

- Pankau, Andreas (2009). Wh-copying in German and its Theoretical Implications. Talk held at the Workshop on Long-Distance Dependencies, Meertens Instituut, Amsterdam, December 2009.
- Rackowski, Andrea and Norvin Richards (2005). Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36(4), 565 - 599.
- Riemsdijk, Henk van (1978). *A case study in syntactic markedness: the binding nature of prepositional phrases*. Lisse: Peter de Ridder Press.
- Schippers, Ankelien. (2009). Some people are repeaters: Medial copy spell-out in long-distance wh-dependencies. To appear in: Proceedings of ConSOLE XVII.
- Schippers, Ankelien and Jack Hoeksema (2009). Variation in long-distance dependencies. Ms., Groningen: Rijksuniversiteit Groningen.
- Thornton, Rosalind (1990). *Adventures in wh-movement*. Ph.D. dissertation, University of Connecticut.
- Verhagen, A (2006). On subjectivity and 'long distance Wh-movement'. In Angeliki Athanasiadou, Costas Canakis and Bert Cornillie (eds.), *Subjectification: Various Paths to Subjectivity*, pp. 323 - 346. Berlin/New York: Mouton de Gruyter.