Syntactic Doubling and the Structure of Chains

Sjef Barbiers\textsuperscript{1,2}, Olaf Koeneman\textsuperscript{1}, and Marika Lekakou\textsuperscript{1}
\textsuperscript{1}Meertens Instituut and \textsuperscript{2}Utrecht University

1. Introduction

A recent survey of 267 Dutch dialects provides five cases of syntactic doubling, in which two elements (WH-pronouns or subject pronouns) referring to the same syntactic entity co-occur in one and the same sentence. This paper focuses on four cases of non-identical doubling, where two distinct-looking elements co-occur. It is shown that the order of these pronouns is fixed and that the first (or syntactically higher) pronoun must be less specific than the second one. It is argued that this generalization follows from 'partial copying', a process that copies a proper subconstituent and remerges it higher in the structure. This naturally excludes the ungrammatical orders, as those would involve full copying plus the addition of features, in violation of the inclusiveness condition. The proposal requires pronouns to be spell outs of phrases (cf. Cardinaletti & Starke 1999 a.o.) and it is in combination with this hypothesis that the full set of data is accounted for in a uniform way. Advantages over alternative accounts of syntactic doubling are briefly discussed.

The paper is structured as follows. Section 2 presents the data that lead to the generalization mentioned above and proposes to derive it from the copy theory of movement. Section 3 argues that pronouns must be spell outs of phrases and offers concrete syntactic structures for them. Section 4 points out interesting parallels of the proposal with the phenomenon of \textit{wat...voor} split. Section 5 shows the advantages over other recent accounts of syntactic doubling. Section 6 concludes.

2. Non-identical doubling and partial copying

The most straightforward cases of syntactic doubling are those that involve identical copies. Example (1) shows doubling of a non-neuter WH-pronoun, and example (2) doubling of a strong subject pronoun:

(1) \textit{Wie} denk je \textit{wie} ik gezien heb? (e.g. Drenthe)
  who think you who I seen have
  ‘Who do you think I have seen?’

(2) \textit{Zij} heeft \textit{zij} daar niets mee te maken. (e.g. Boutsersen)
  she has she there nothing with to do
  ‘She has got nothing to do with it.’

In addition, Dutch dialects show patterns of non-identical doubling. In example (3), a neuter WH-pronoun co-occurs with a non-neuter WH-pronoun. In example (4), a non-neuter WH-pronoun co-occurs with a relative pronoun. Example (5) combines a neuter WH-pronoun with a relative pronoun. An example of non-identical subject doubling is given in (6), where a weak and a strong pronoun co-occur.

(3) \textit{Wat} denk je \textit{wie} ik gezien heb? (e.g. Overijssel)
  what think you who I seen have

(4) \textit{Wie} denk je \textit{die} ik gezien heb? (e.g. North-Holland)
  who think you rel.pron I seen have

Wat denk je die ik gezien heb? (e.g., Overijssel)
what think you rel.pron I seen have

Ze heeft zij daar niks mee te maken. (e.g. Flemish)
she.weak has she.strong there nothing with to do
'She’s got nothing to do with it.'

The four cases of distinct doubling have in common that the reverse order of elements is strongly ungrammatical.

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

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

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

*Zij heeft ze daar niks mee te maken.
she.strong has she.weak there nothing with to do

One could object that (8) and (9) are out for the independent reason that relative pronouns cannot introduce a WH-question, but note that in relative clause constructions reversing the order is equally impossible:

a. Dit is de man wie ik denk die Jan gezien heeft. (e.g. Drenthe)
this is the man who I think rel.pron I have seen.'
b. *Dit is de man die ik denk wie Jan gezien heeft.
this is the man rel.pron I think who Jan seen has
'This is the man Jan thinks I have seen.'

The central question is whether the two pronouns in doubling constructions are part of the same syntactic chain. It seems natural to assume for (1) and (2) that the pronouns have undergone syntactic movement and that at PF more than one chain link is spelled out. We would therefore like to pursue the default hypothesis that all the examples presented involve syntactic chains, also the cases of non-identical doubling. Under this assumption, the following generalization can be formulated (cf. Barbiers 2006):

In a syntactic movement chain, a higher chain link cannot be more specified than a lower chain link.

Let us illustrate this using example (6). One could say that the weak pronoun, ze, is specified as [3rd person, singular, female]. Strong pronoun zij is also marked for these features, but in addition carries a focus feature: [3rd person, singular, female, focus]. Hence, (12) correctly predicts that ze must be higher in the structure than zij and will therefore always precede it.

The next step is then to explain what causes (12) to hold. We propose that it follows from the copy theory of movement (Chomsky 1995). This theory allows the syntax to copy a constituent α and remerge α higher in the structure. This will give standard movement. What syntax should also be

---

1 For examples like (3) in German, it has been argued that a syntactic chain is not created until LF, where the lower WH-pronoun moves to replace the higher one (cf. van Riemsdijk 1982; McDaniel 1989; Beck & Berman 2000). Others have argued that there is no direct dependency at all and that the higher WH-pronoun is an argument of the matrix predicate (cf. Dayal 1994; Felser 2001). Although we have no space to show this, we believe that the arguments for the latter position do not easily carry over to Dutch. Hence, we will pursue the default hypothesis.
allowed to do is to partially copy α. This is what happens with subextraction: trivially, an object must be able to move out of the VP, stranding the rest of the VP. Hence, (13) summarizes the options available to the grammar:

(13)

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Phonology</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Full copying</td>
<td>(a) Spell out one chain link</td>
</tr>
<tr>
<td>(b) Partial copying</td>
<td>(b) Spell out more than one chain link</td>
</tr>
</tbody>
</table>

Suppose now that we apply these options to the issue at hand and assume that copying of a pronoun can also be either full or partial (cf. Cheng 2000, Sabel 2000). This gives the following possible outputs:

(i) Full copying, but (for some reason) only the higher chain member is spelled out: non-doubling.
(ii) Full copying and both chain members are spelled out: identical doubling.
(iii) Partial copying and both chain members are spelled out: non-identical doubling.

What is now naturally excluded is a case in which the more specified pronoun precedes the less specified one. In order to generate such a construction, one would have to make a full copy of a pronoun and subsequently add features and/or structure to it. This addition of features does not follow from the copy theory. It in fact violates the inclusiveness condition (Chomsky 1995), which states that the output of a syntactic operation cannot contain anything beyond its input.

3. Partial copying and the phrasal analysis of pronouns

The proposal in the previous section raises an immediate question. What does it mean to partially copy a pronoun like wie? What it cannot mean is that wie is a head, the terminal string in a syntactic representation, and that partial copying targets a subset of its features. This would essentially entail that a syntactic operation, partial copying, is allowed to apply to elements at a sub-word level, violating lexical integrity (Lapointe 1980). Adopting partial copying as the mechanism responsible for non-identical doubling essentially forces one to assume (like Cardinaletti & Starke 1999; Déchaine & Wiltschko 2002; van Koppen 2005; Neeleman & Szendröi 2006, a.o.) that pronouns are not spell outs of terminals but spell outs of phrases. Adopting this view, we will argue that die, wie and wat spell out different layers of the nominal projection, as indicated in (14):

\[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{PhiP} \\
\text{Phi} \\
\text{QP} \\
\end{array} = \begin{array}{c}
die \\
wie \\
wat \\
\end{array}
\]

Given this structure, partial copying can now target the PhiP- or QP-node and the result will be non-identical doubling. Let us first discuss these nominal layers in turn (section 3.1) and then provide some derivations of doubling constructions (section 3.2).

3.1 The analysis of wat, wie and die

Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).

\[\text{Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).}\]

\[\text{Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).}\]

\[\text{Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).}\]

\[\text{Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).}\]

\[\text{Wat can be many things in Dutch. It can be an indefinite argument (15a), a nominal modifier (15b), a relative pronoun (15c), a wh-pronoun (15d) or an exclamative marker (15e).}\]
Instead of listing \textit{wat} in the lexicon six times, we follow Postma (1994) and assume that there is only one \textit{wat}, its interpretation being determined by its syntactic context (cf. Cheng 1991 for a similar claim for Chinese indefinites). It is shown in (16) that in object position \textit{wat} is interpreted as an indefinite, whereas in clause-initial position it is interpreted as a WH-pronoun.

\begin{enumerate}
  \item Jan heeft \textit{wat} gegeten.
  \begin{itemize}
    \item Jan has \textit{wat} eaten
    \begin{itemize}
      \item Jan has eaten something.'
    \end{itemize}
    \item *What has Jan eaten?'
  \end{itemize}
  \item Wat heeft Jan gegeten.
  \begin{itemize}
    \item What shall I eat tonight?'
  \end{itemize}
  \item Wat een ellende is dit!
  \begin{itemize}
    \item 'What a disaster this is!'\end{itemize}
\end{enumerate}

For \textit{wat} to appear in this variety of contexts, it must be rather unspecified itself. This indeed seems to be the case. \textit{Wat} can modify neuter and non-neuter non-count nouns (17), suggesting it is unspecified for gender. It can also modify a plural noun (18), suggesting it is not specified for number. In addition, \textit{wat} can appear in an expletive construction, which in Dutch triggers a strong definiteness effect. Hence, the grammaticality of (19) suggests that \textit{wat} is not specified for definiteness either:

\begin{enumerate}
  \item Jan heeft \textit{wat} gegeten.
  \begin{itemize}
    \item Jan has \textit{wat} eaten
    \begin{itemize}
      \item Jan has eaten something.'
    \end{itemize}
    \item *What has Jan eaten?'
  \end{itemize}
  \item Wat heeft Jan gegeten.
  \begin{itemize}
    \item What shall I eat tonight?'
  \end{itemize}
  \item Wat een ellende is dit!
  \begin{itemize}
    \item 'What a disaster this is!'\end{itemize}
\end{enumerate}

\begin{enumerate}
  \item het brood / \textit{wat} brood
  \item the neuter bread \textit{wat} bread
  \item de kaas / \textit{wat} kaas
  \item the non-neuter cheese \textit{wat} cheese
\end{enumerate}

\begin{enumerate}
  \item wat boeken
  \begin{itemize}
    \item \textit{Wat} book plural
    \begin{itemize}
      \item 'some books'
    \end{itemize}
  \end{itemize}
  \item Er is \textit{wat} gekomen.
  \begin{itemize}
    \item there is \textit{wat} come
    \begin{itemize}
      \item 'Something arrived.'\end{itemize}
  \end{itemize}
\end{enumerate}

As indicated in (20), we think \textit{wat} is the spell out of a QP. This makes \textit{wat} close to Dutch \textit{veel} 'many' and \textit{weinig} 'little'. \textit{Veel} and \textit{weinig} have been analyzed as Q-elements in the literature (cf. Corver 1997). The denotations we informally assign to them are given in (20):

\begin{enumerate}
  \item \textit{veel} = high quantity; \textit{weinig} = low quantity; \textit{wat} = quantity
\end{enumerate}
There are three arguments in favour of the suggested parallel. First, all three can be modified by adverbs like heel 'very' and nogal/best 'quite':

\[
21 \begin{cases}
\begin{align*}
\text{heel} & \text{veel/weinig/wat boeken} \\
\text{nogal} & \text{veel/weinig/wat boeken} \\
\text{best} & \text{veel/weinig/wat boeken}
\end{align*}
\end{cases}
\]

Secondly, they appear in complementary distribution with elements such as determiners and numerals.

\[
22 \begin{align*}
a. & \text{de veel/weinig/wat boeken} \\
& \text{the many/few/some books}
\end{align*}
\]

\[
22 \begin{align*}
b. & \text{twee veel/weinig/wat boeken} \\
& \text{two many/few/some books}
\end{align*}
\]

Thirdly, they can all co-occur with quantitative er, which is a property of numerals:

\[
23 \begin{align*}
\text{Ik heb er veel/weinig/wat boeken.} \\
\text{I have ER many/few/WAT about read} \\
'\text{I read much/little/something about it.}'
\end{align*}
\]

For the analysis of wie, we following Dechaine and Wiltischko (2002) in assuming that DPs can have a PhiP-layer, where among other things gender is expressed. It is this layer that can be spelled out by wie (cf. 24a). The difference between wie and die then amounts to the presence or absence of a DP-layer (cf. 24b).

\[
24 \begin{align*}
a. \text{PhiP} = \text{wie} \\
\text{Phi' QP} = \text{wat} \\
\text{[+gender]}
\end{align*}
\]

\[
24 \begin{align*}
b. \text{DP} = \text{die} \\
\text{D' PhiP} = \text{wat}
\end{align*}
\]

The contribution of the DP-layer is definiteness (Bennis 2001), as indicated in (25):

\[
25 \begin{align*}
a. & \text{Dat = wat} + \text{definiteness} \\
b. & \text{Die = wie} + \text{definiteness}
\end{align*}
\]

We have already seen that wat cannot appear in a topic position without triggering a question interpretation (cf. 26a). This is in contrast to dat (cf. 26b), which suggests that this element is definite.

\[
26 \begin{align*}
a. & \text{Hans heeft wat gelezen.} \quad \Rightarrow \text{Wat heeft Hans gelezen?} \\
& \text{Hans has WAT read} \quad \text{WAT has Hans read} \\
& '\text{Hans has read something.'} \quad *'\text{Something, Hans has read.'}
\end{align*}
\]

\[
26 \begin{align*}
b. & \text{Hans heeft dat gelezen.} \quad \Rightarrow \text{Dat heeft Hans gelezen.} \\
& \text{Hans has that read} \quad \text{that has Hans read} \\
& '\text{Hans has read that'.} \quad '\text{That, Hans has read.'}
\end{align*}
\]

In virtue of being able to function as topics, dat and die can be topic-dropped, again unlike wat and wie:

\[
27 \begin{align*}
a. & \text{Wat je weet, (dat) mag je zeggen.} \\
& \text{WAT you know D-pron. may you say} \\
& '\text{Whatever you know you can tell.}'
\end{align*}
\]

\[
27 \begin{align*}
b. & \text{Wie het weet, (die) mag het zeggen.} \\
& \text{who it knows D-PRON. may it say-INF.} \\
& '\text{Whoever knows it may say it.'}
\end{align*}
\]
We assume that *wie* has no WH-feature, as that would imply that *die* has one too. The lack of a WH-feature is suggested by the fact that both *wie* and *die* can function outside the context of WH-questions, namely as relative operators.

To summarize, *die, wie* and *wat* are spell outs of phrases corresponding to different layers of a nominal structure, DP, PhiP and QP, respectively. With this in place, we can now go over a few illustrative examples of doubling constructions. Before we do so, we will make one additional assumption, namely that PhiP- and DP-internally, the QP moves to the highest specifier position, as indicated for PhiP in (28a). This allows us to draw an interesting parallel with *wat...voor*-split, a construction that also allows subextraction of *wat* (cf. 28b). We will discuss this parallel in more detail in section 4.

3.2 Deriving the generalization

The examples of doubling in section 2 can now be derived by ensuring that minimally the QP is copied. The different patterns arise because of (i) the optionality of generating either a DP or PhiP in the lower position and (ii) the optionality of pied-piping DP and PhiP. We will not go over all the examples but provide two examples of full copying, two of partial copying and one of an ungrammatical doubling construction, that is one that violates the generalization in (12).

The most straightforward case is that of full copying. This is for instance derived by (i) copying a PhiP, (ii) remerging it higher in the structure and (iii) spelling out two chain links:

(28) a. \[ \text{PhiP} = \text{wie} \]

b. \[ \text{XP} \]

(29)

\[
\begin{array}{c}
\text{PhiP} \\
\text{Phi+gender}[\text{QP} \ldots] \\
\end{array}
\]

\[
\begin{array}{c}
\text{PhiP} \\
\text{Phi+gender}[\text{QP} \ldots] \\
\end{array}
\]

cf. (1) \begin{tabular}{ll}
Wie & denk je \\
who & think you
\end{tabular}

\begin{tabular}{ll}
wie & ik gezien heb? \\
who & I seen have
\end{tabular}

Identical subject doubling (cf. example (2)) is derived by fully copying a strong pronoun, which we analyze as a FocP:

(30)

\[
\begin{array}{c}
\text{Foc} \\
\text{Foc} \quad \text{[phiP \ldots]} \\
\end{array}
\]

\[
\begin{array}{c}
\text{Foc} \\
\text{Foc} \quad \text{[phiP \ldots]} \\
\end{array}
\]

cf. (2) \begin{tabular}{ll}
Zij & heeft \\
she.weak & has
\end{tabular}

\begin{tabular}{ll}
\text{zij} & daar niks mee te maken. \\
she.strong & there nothing with to do
\end{tabular}

Partial copying is derived by copying only the PhiP (cf. 31) or the QP (cf. 32):

(31)

\[
\begin{array}{c}
\text{PhiP} \\
\text{Phi+gender}[\text{QP} \ldots] \\
\end{array}
\]

\[
\begin{array}{c}
\text{DP} \\
\text{D+definite \quad [phiP \text{Phi+gender}[\text{QP} \ldots]]} \\
\end{array}
\]

cf. (4) \begin{tabular}{ll}
wie & denk je \\
who & think you
\end{tabular}

\begin{tabular}{ll}
die & ik gezien heb? \\
rel.pron & I seen have
\end{tabular}
Finally, example (33) illustrates what goes wrong in an ungrammatical doubling case. To derive an example where the first element is more specified than the second, full copying has to be followed by something that the copy procedure itself does not provide, namely the addition of features and/or structure:

4. The parallel with \textit{wat...voor}-split

A crucial ingredient in the analysis is that partial copying should be allowed to target a specifier which is part of a constituent in a derived position. Although this may not seem the most straightforward position from which to allow extraction (but see Chomsky 1986: p. 25 ff.), this is exactly what seems to happen in the \textit{wat...voor}-split (cf. 34b):

There are in fact four additional similarities between partial \textit{WH}-copying and \textit{wat...voor}-split.

First of all, the higher copy cannot be more specified than the lower one, in line with (12).

Secondly, PF never spells out \textit{WH}-copies VP-internally.

Not spelling out the lower \textit{wat} in (36b) gives a grammatical result (cf. 37b). Not spelling out the lower copy is not a valid strategy for (36a), as that would delete unrecoverable Phi-features, i.e. the sentence can no longer get the \textit{who}-reading (cf. also footnote 2).
a. Ik vraag me af [CP wat Jan wie gezien heeft].
   I ask me PRT what John who seen has
   'I wonder what Jan has seen.'

b. Ik vraag me af [CP wat Jan wat voor boeken gelezen heeft].
   I ask me PRT what Jan for books read has
   'I wonder what kind of books Jan has read.'

Thirdly, when they occur in SpecCP of the embedded clause, WH-elements must be spelled out:

(38) a. [CP wat denk je [CP *(wie) Jan gezien heeft?]]
   what think you who John seen has
   'Who do you think Jan saw?'

b. [CP *wat denk je [CP *(wat) voor boeken Jan gelezen heeft?]]
   what think you what for books Jan read has
   'What kind of books do you think that Jan has read?'

A last similarity is that both partial WH-copying and 'wat...voor'-split are blocked by negation (cf. Felser 2004 and references cited therein):

(39) a. *Wat denk je niet wie Jan ontmoet heeft?
   what think you not who John met has
   'Who do you think that Jan met?'

b. *Wat denk je niet dat Jan voor mensen ontmoet heeft?
   what think you not that John for people met has
   'What kind of people do you think that Jan met?'

These similarities strengthen the idea that partial WH-doubling and wat...voor-split are essentially derived by the same process which only partially copies a constituent. The main difference is that in wat...voor-constructions wat is always visible, whereas in WH-pronoun constructions wat only shows up if partial copying takes place. When full copying applies, or no copying at all (i.e. there is no movement), wat is not spelled out, because spell out of wie subsumes spell out of wat. In wat...voor-constructions, on the other hand, voor as well as the noun are spell outs of terminal nodes, neither of which dominates wat. Hence, wat can and will be spelled out.

5. Alternative approaches

Recent literature offers two alternative approaches to partial doubling. Nunes (2004) suggests to treat the phenomenon in terms of scattered deletion. Poletto & Pollock (2004) have argued for an approach in which doubler and doublee initially form one big XP. We will discuss these proposals in turn and note why we think that, at least for the data under discussion, our own approach is to be favoured.

Nunes argues that spelling out more than one chain member leads to a violation of Kayne's LCA. If syntactic structure is linearized on the basis of c-command relations, movement of α across β leads to the situation in which β both c-commands and is c-commanded by α. Hence, α and β cannot be linearized. One strategy to allow linearization is to delete all but one copy. Another strategy is to delete complementary material in two copies. (40) provides an example from Croatian:

(40) a. Na kakav je Ivan krov bacio loptu? (Croatian)
   On what.kind.of be Ivan roof throw ball
   'On what kind of roof did Ivan throw the ball?'

b. ([PP na [kakav krov]]) je Ivan ([PP na [kakav krov]]) bacio loptu ([PP na [kakav krov]])

There is no partial copying under this account. In the syntax, the whole WH-constituent is copied and PF subsequently decides where to delete what (cf. 40), thereby deriving what looks like subextraction. As a tool for the data under discussion, however, scattered deletion is not precise enough as it stands. It wrongly predicts partial underspecification of a lower copy to be possible, so additional stipulations are needed to derive the generalization in (12).
In their analysis of WH-clitic doubling in French and Northern Italian dialects, Poletto & Pollock argue that doubler and doublee are initially contained in one and the same nominal XP. For the data at hand, these XPs would look like the ones given in (41):

(41) a. [XP wat wie] or [XP wie wat]
    b. [XP wie die] or [XP die wie]
    c. [XP ze zij] or [XP zij ze]

During the derivation one of them moves out. There are two reasons for not adopting this big XP analysis. First of all, the XPs in (41) never occur overtly as one constituent, but only discontinuously. This is unexpected if subextraction of WH from big XPs is parallel to wat-voor split (cf. Sabel 2000), since pied-piping of the voor-NP is a general option (cf. 34a). Second, the big XP analysis requires additional stipulations to block extraction of the more specified element: Moving wie out of [XP wat wie] would lead to a construction that violates (12). In the present analysis, neither problem arises. The more specified element in a doubling chain is the spell-out of a category (PhiP/DP) that dominates the category that would be spelled out by the less specified element. Hence, the less specified element only becomes visible when it moves out of the PhiP/DP.

6. Conclusion

In this paper, we have argued that copying can be full or partial. In combination with a phrasal analysis of pronouns, we are able to correctly distinguish between the grammatical cases of WH- and pronoun doubling and the ungrammatical ones. In addition, we have shown why alternative analyses of doubling cannot account for the data under discussion.

The generalization in (12), which served as the input to the analysis, makes a strong claim and there are of course counterexamples to it. Contrastive left dislocation in Germanic, clitic left dislocation in Romance, and constructions that involve resumptive pronouns have in common that the higher element is more specified than the lower one. Before we conclude that these constructions falsify (12), however, it has to be established beyond a fair amount of doubt that, here too, the higher and lower element are part of the same syntactic chain. The task of future research is therefore to embed the current analysis into this broader discussion.

References


