Prepositional numerals and spurious determiners

A.H.J. (Astrid) van Alem
Leiden University Centre for Linguistics, Leiden University, Leiden, Netherlands.
Manuscript written during their RMA Linguistics, Utrecht University, Utrecht, Netherlands.

Abstract
This paper deals with the syntax of prepositional numerals, such as around 20 people. English contrasts with Dutch, in that Dutch requires presence of a definite article (tegen *(de) 20 mensen “against the 20 people”) whereas in English this is ungrammatical (“around the 20 people”). This paper argues that the article in Dutch is not a definite determiner. Instead, it is proposed to combine Corver and Zwarts (2006) and Barbiers (2005), arguing that the numeral PP merges as the complement of the noun that it modifies, and that within this PP there is an empty noun that needs to be licensed by valuing its only unvalued feature via Reverse Agree (Wurmbrand 2014). In Dutch this feature is unval[gender] leading to the spell out of a morpheme with val[gender], namely de. In English, this feature is unval[count], which can receive its v under agreement with the numeral, and thus no additional morpheme needs to be inserted. Under such an analysis, many of the exceptional properties of the article in prepositional numerals in Dutch follow. This is in line with the idea that syntactic variation can be reduced to variation in the lexicon (Chomsky, 1995).

1. Introduction
Languages use different ways of expressing numeral approximation. One way of doing so is by the use of prepositional phrases, as illustrated in (1) and (2) for Dutch and English. 2

(1) a. Ik las [pp in de 20 boeken].
I read in the 20 books
"I read 20-something books."

b. Ik las [pp tegen de 20 boeken].
I read against the 20 books
"I read nearly 20 books."

(2) a. Ik las [pp around 20 books].
I read around 20 books.
b. Ik las [pp between 15 and 25 books].

1 I thank Marijana Marelj, whose comments significantly improved this paper, as well as three anonymous LingUU reviewers for helpful suggestions.
2 All non-English examples in this paper are in Standard Dutch.
Corver and Zwarts (2006) investigated the syntax and semantics of these prepositional numerals, finding interesting parallels between the clausal and the nominal domain in terms of displacement. However, there are still many properties unexplained about these constructions. For instance, it remains unknown why in Dutch prepositional numerals typically the article *de* “the” is used, whereas it is not present in the English counterparts. It is this phenomenon that I will investigate in this paper.

The structure of the paper is as follows. Section 2 shows that the view of the article as a determiner for the modified noun cannot be maintained. Instead, section 3 proposes an analysis in which the numeral PP merges as the complement of the noun that it modifies (Corver & Zwarts, 2006). I propose that in this numeral PP, there is an empty noun, and that the empty noun needs to be licensed by expressing the feature [gender] in Dutch (in line with Barbiers 2005). This feature is morphologically realized as the article in the numeral PP. In section 4, I will get into more detail of the theory and many properties of the Dutch and English prepositional numerals follow. Section 5 concludes with some open questions and suggestions for future research.

2. What the article is not
The article *de* “the” in Dutch prepositional numerals is superficially similar to the determiner in ‘ordinary’ PPs (3, 4).

(3) *Marie heeft tegen de 8 boeken gelezen.*
Mary has against the 8 books read
“Mary has read nearly 8 books.”

(4) *De gasten zitten tegen de tafel.*
the guests sit around the table
“The guests are sitting against the table.”

The ‘naive’ view would be that the constructions are in fact identical, and that the article *de* is a definite determiner for the noun *boeken* “books”. However, on deeper inspection the constructions appear to have different properties, suggesting a different analysis.

First of all, if the article in the prepositional numeral were of the same nature as the article in ordinary PPs, one would expect it to have the same distribution among the constructions. However, this is not the case, as illustrated in (5) and (6).

(5) a. *De gasten zitten tegen de 8 tafels.*
the guests sit against the 8 tables
“The guests are sitting against the 8 tables.”

b. *De gasten zitten tegen 8 tafels.*
the guests sit against 8 tables
“The guests are sitting against 8 tables.”
In the room stand against the 8 tables.

"In the room there are nearly 8 tables."

In the room stand against 8 tables
intended: "In the room there are nearly 8 tables."

Although it is fine to leave out the determiner in the normal PP (5b), giving rise to a non-referential reading for *tafels* "tables", leaving out the determiner in the prepositional numeral (6b) leads to an ill-formed sentence. The same holds for other prepositional numerals, e.g. numerals PP containing the preposition *in* (7):

Jan heeft in 20 mensen uitgenodigd.
Jan has in 20 people invited
intended: "Jan invited more than 20 people."

Furthermore, the use of the definite article in the prepositional numeral does not have the semantic effect that definite determiners normally have. In its normal use, the definite determiner combined with a noun or quantified noun will lead to a specific interpretation (i.e. the speaker refers to an already identified set of individual). In prepositional numerals this is not the case. In (3), repeated (8), the books that Mary read can be any set of books with a cardinality of nearly 8. Hence, the definite article does not make the phrase specific.

Marie heeft tegen de 8 boeken gelezen.
Mary has against the 8 books read
"Mary has read nearly 8 books."

The observation that the definite article in prepositional numerals does not make the whole phrase definite can also be illustrated syntactically, by means of the definiteness restriction. The definiteness restriction is the generalization that only indefinite noun phrases can occur as the subject of sentence beginning with expletive *there* (Milsark, 1974). Prepositional numerals are perfectly fine as the subject of such a sentence, as illustrated in (9a, cf. 9b). This shows that the construction is syntactically indefinite, despite the presence of the article.

The effect is somewhat less strong with e.g. *rond* "around", as (i) illustrates:

(i) a. Ik heb rond de 20 mensen uitgenodigd.
I have around 20 people invited
"I have invited around 20 people."

b. ?Ik heb rond 20 mensen uitgenodigd.
I have around 20 people invited
"I invited around 20 people."

I will discuss these cases in section 4.
(9)  
\begin{align*}
&\text{a. } \text{Er liggen tegen de 8 boeken op tafel.} \\
&\quad \text{there lie against the 8 books on table} \\
&\quad \text{“There are nearly 8 books on the table.”}
\end{align*}

\begin{align*}
&\text{b. } \text{Er liggen de 8 boeken op tafel.} \\
&\quad \text{there lie the 8 books on table} \\
&\quad \text{intended: “There are the 8 books on the table.”}
\end{align*}

Another piece of evidence against the view that the article in prepositional numerals is an ordinary determiner belonging to the noun is the following. Dutch makes a distinction between neuter and common gender, and this is visible on the definite determiner (and in prenominal adjectival inflection): neuter words require the determiner *het*, whereas common words require *de*. However, when a neuter nominal like *jaar* “year” is part of a prepositional numeral, the common article *de* is obligatorily used (10, cf. the contrast with 11), indicating that the determiner is not in an Agreement relation with the noun.\(^4\)

\begin{align*}
&\text{(10) } \text{Rond de 1 jaar.} \\
&\quad \text{around the 1 year} \\
&\quad \text{“Around 1 year.”}
\end{align*}

\begin{align*}
&\text{(11) } \text{Het jaar.} \\
&\quad \text{the year} \\
&\quad \text{“The year.”}
\end{align*}

Furthermore, it is in general ungrammatical to use a determiner in combination with the numeral 1 in a non-prepositional numeral context (12). This is a further indication that the structure inside the prepositional numeral is not that of an ordinary nominal phrase, but has different properties.\(^5\)

\begin{align*}
&\text{(12) } \text{Het 1 jaar.} \\
&\quad \text{the 1 year} \\
&\quad \text{“The one year.”}
\end{align*}

Concluding, in this section I have argued that the article in prepositional numerals is not an ordinary determiner, despite superficial similarities. This makes it unlikely that the underlying structures are identical. In the next section, I will propose a structural analysis of the prepositional numeral construction.

3. Analysis

3.1. Theoretical background

In order to understand the properties of numeral PPs, it is necessary to be explicit about

\(^4\) The prepositional numeral construction with “1” appears to be limited to nouns that denote a timespan, e.g. day, week, year, etc. This might be due to the fact that it is possible to have parts of those concepts, whereas this is not straightforwardly possible with concepts like ‘house’ or ‘cat’.

\(^5\) Note that when 1 (één) inflects (één-e), (12) is grammatical. I will come back to this in section 4.
the structure of the PP and the modified DP. This paper works with the articulated structure of DP in (13) (Ritter, 1991).  

In this structure nouns merge in N°, determiners in D°. Following Corver and Zwarts (2006), I assume that numerals are base-generated as the complement of N°, and move to Spec,NumP as an instance of DP-internal predicate movement. The structure of a DP containing a numeral PP (around 8 books) is then as in (14) (strike indicates a copy).

\[(13)\quad \text{DP} \quad (14)\quad \text{DP} \]

This analysis implies that numerals are not heads (contra Giusti, 1991; Barbiers, 2005), and that they merge to the right of the noun. An argument for the former is that numerals can be phrasal (e.g. complex cardinals or prepositional numerals), and thus cannot be heads (Corver & Zwarts, 2006). With respect to the latter, Corver and Zwarts show that in all other syntactic environments, PPs merge postnominally. By analogy, the numeral preposition is expected to also merge to the right of N°. Indeed, there are cases in which a numeral PP occurs postnominally (albeit with a different meaning – see Corver & Zwarts for further discussion) (15, from Corver & Zwarts). This can be easily captured with the analysis that assumes base-generation of the numeral PP as the complement of N°.

\[(15)\quad \text{Er staan } \left[\text{PP getallen }\right] \text{ op het bord.} \]
\[\text{there stand } \text{numbers above }100 \text{ on the blackboard} \]
\[\text{"there are numbers above 100 on the blackboard"} \]

3.2. The internal structure of numeral PP and syntactic variation

In the previous section I have adopted the structure of prepositional numerals as proposed in Corver and Zwarts (2006). According to this structure, the numeral PP (around 8) merges as the complement of the noun that it modifies (books). What, then, are the internal properties of the numeral PP?

---

6 Note that it is likely that there are additional projections in the DP, such as a PhiP (cf. Déchaine & Wiltchko 2002) or focus projections (cf. Aboh 2004). However, the focus of this paper is on the relation between numerals and nouns, therefore the structure in (13), where the NumP is explicitly present, will suffice for the present purposes.
I would like to propose that inside the PP there is another DP with the numeral in Spec, NumP. The DP lacks an overt NP; instead I propose that \( N^\circ \) is filled with the empty noun \( e \) (in the spirit of Panagiotidis, 2003). The internal structure of the PP is given in (16). The full structure of the prepositional numeral is given in (17).

The nominal element \( e \) is devoid of semantic and phonological features, but it needs to be licensed syntactically by realizing a syntactic feature elsewhere in the structure.\(^7\) I propose this goes against Panagiotidus’ original proposal, but see Barbiers (2005) fn. 6 for arguments that a licensing mechanism is needed in Panagiotidis’ theory nevertheless.
to formalize this as an unvalued feature on $e$ that needs to be in Agreement with a corresponding valued feature via Reverse Agree as in the definition in (18) (Wurmbrand, 2014).

(18) Reverse Agree:
A feature $\text{unval}[F]$ on $\alpha$ is valued by a feature $\text{val}[F]$ on $\beta$, iff
i. $\beta$ c-commands $\alpha$ AND
ii. $\alpha$ is accessible to $\beta$ (accessible: not spelled-out).
iii. $\alpha$ does not value a feature of $\beta/a$ feature $F$ of $\beta$.

In other words: the unvalued feature on $e$ can be valued by an element bearing a corresponding valued feature that has $e$ in its c-command domain (requirement i) within (at least) the PP (requirement ii). Requirement iii does not play a role here, since $e$ is devoid of all features beside its licensing feature.

I assume that the structure of numeral PPs as in (16) is universal. However, following Barbiers (2005) I propose that languages differ with respect to the nature of the unvalued feature on $e$, giving rise to morphosyntactic variation across languages. Before looking at prepositional numerals, let us see how this works in Barbiers’ proposal for noun ellipsis. Barbiers argues that in noun ellipsis, $e$ is inserted in the place of the elided noun. Based on the contrast in (19) and (20), he argues that in Dutch, the empty noun is licensed by expressing [gender], and in English, the empty noun is licensed by [count].

(19) (Over boeken gesproken,) ik heb een leuk-e gelezen.
(talking about books,) I have a nice-INFL read
“Talking about books, I have read a nice one.”

(20) Talking about books, I have read a nice one.

In the Dutch example in (19), the adjective inflects for [gender] to license noun ellipsis. In my terms, the gender inflection is necessary to value $\text{unval}[\text{gender}]$ on $e$. In (20), the numeral $\text{one}$, bearing [count] is inserted to license noun ellipsis. In my terms: $\text{one}$, bearing a $\text{val}[\text{count}]$ feature, is inserted to value $\text{unval}[\text{count}]$ on $e$. I propose that the same licensing requirements hold in prepositional numerals, leading to the licensing conditions in (21):

(21) License $e$ by valuing its $\text{unval}[\text{feature}]$ via Reverse Agree.
   English: $\text{unval}[\text{count}]$ Dutch: $\text{unval}[\text{gender}]$

---

8 At a minimum, the PP is part of the spell-out domain of $e$, since DP-internal predicate movement of the PP takes place before spell-out. Further research is necessary to determine whether the spell-out domain extends outside of the PP.

9 Note that the adjective always inflects with the schwa, even if it would not if the noun was not elided (i.e. when the noun is a singular neuter). This is consistent with an empty noun that is maximally underspecified (except for its licensing feature) and morphemes that are either specified for specific feature values or as default or ‘elsewhere’ (as in Halle & Marantz’s(1993) Distributed Morphology framework): schwa-inflection on Dutch adjectives is the least specific (elsewhere) form of inflection and therefore this inflection will be inserted in contexts where there is no indication to insert the more specific morpheme, for instance in the case of an underspecified empty noun.
Let us see how this proposal accounts for the observed facts from English and Dutch. In English, the empty noun in N° is licensed by valuing unval[count]. It is valid to assume that numerals are valued for [count]. For instance, numerals can only occur with count nouns and not with mass nouns (cf. two books; *two money). In the structure of the prepositional numeral as proposed in (16), the numeral has e in its c-command domain within the PP and is therefore able to value unval[count] on e. This leads to a numeral PP without additional morphology.

The case of Dutch is somewhat less straightforward. In Dutch, the empty noun is licensed by valuing its unvalued [gender] feature. In the structure in (16), no element valued for [gender] is present. Hence, in order for the derivation not to crash, an element with a value for [gender] needs to be inserted. I propose that for this reason, the article de is inserted in D°. In Dutch, articles agree in gender (neuter or non-neuter) with noun phrases, and thus we can reasonably assume that articles have valued [gender]. De has e in its c-command domain within the PP and can therefore Agree with e and value its unval[gender] via this operation. This leads to a numeral PP with an (apparent) determiner de. In other words, de is only present because of the licensing requirement of realizing [gender].

One might wonder why it is always the article de that is inserted to value unval[gender] on e in Dutch, and not het. This is straightforwardly explained by the nature of the empty nominal and the feature specifications of Dutch definite determiners. In Dutch, the determiner het appears to be the marked variant compared to de: het is only used with singular neuter nouns, where de is used in all other cases. In terms of a Distributed Morphology framework (Halle & Marantz, 1993), de would be specified as the elsewhere morpheme. I have argued that the empty nominal e is underspecified for all formal features except its licensing feature (in Dutch: unval[gender]). Hence, there is no indication to insert the more specific morpheme het and therefore the elsewhere article de will always be inserted.

4. Predictions, support, and challenges

The analysis proposed in the previous section accounts for the main syntactic difference between Dutch and English prepositional numerals, namely the obligatory presence of a definite article in Dutch, but not in English. I proposed that the article is present because of licensing requirements. In this section I will discuss some remaining questions and show that the analysis accounts for most of the other properties of prepositional numerals as well.

In section 2 we saw that the article in Dutch prepositional numerals has some unexpected properties compared to ‘real’ determiners. These properties follow straightforwardly from the analysis presented above. First, we saw that the article can mismatch in gender with the noun that is modified by the prepositional phrase in Dutch. This follows from the account presented above: the article is inserted to license unval[gender] on e within the PP. Because e is underspecified, it will always be the elsewhere (common) article de that is inserted. The gender of the noun that the numeral PP modifies is independent of the valuation operation. Hence in the case where the modified noun is neuter, an apparent gender mismatch between the article and the noun will arise.
A related property is that in normal DPs, a determiner in combination with the numeral 1 is ungrammatical, whereas it is grammatical in prepositional numerals. This also follows if we assume that the article in the prepositional numeral is part of the PP, rather than in a projection directly above the noun: it is not affected by the noun's gender.

I have argued that de needs to be present in Dutch numeral PPs in order to license the empty nominal. The question arises if also outside the context of numeral PPs de can be an empty noun licenser. This appears to be the case, as illustrated by the following. In Dutch, a numeral on its own is not sufficient as an empty noun licenser (in contrast to English) (22a). In order to make this possible, some element with val[gender] has to be inserted. Most commonly, this will be er “there” (22b), from which we can conclude that er has a val[gender] feature. However, in specific semantic contexts, de can also be used with the same effect on licensing (22c). In this case, it is ungrammatical to use er as well, indicating that de is able to license empty nouns (22d).

Heb je boeken over taalwetenschap? “Do you have books about linguistics?”

(22) a. Ja, ik heb twee van Chomsky.
   yes I have two of Chomsky
   intended: “Yes, I have two (books) by Chomsky.”

b. Ja, ik heb er twee van Chomsky.
   yes I have there two of Chomsky
   “Yes, I have two (books) by Chomsky.”

c. Ja, ik heb de twee van Chomsky.
   yes I have the two of Chomsky
   “Yes, I have the two (books) by Chomsky.”
   (Interpretation: Chomsky has written only 2 books on linguistics and this is common knowledge.)

d. Ja, ik heb er de twee van Chomsky.
   yes I have there the two of Chomsky
   intended: “Yes, I have the two (books) by Chomsky.”

One might ask why er cannot be used inside numeral PPs to value unval[gender] on e (“rond er 8 boeken ”around there 8 books”). It appears that in order for er to occur in a PP, it has to be able to move out of the PP to some higher position in the clause (23). 10

(23) a. Ik heb ook op de winst gerekend.
   I have also on the victory counted
   “I also counted on the victory.”

10 In the examples the adverb ook is included to show that er does not move to Spec,PP but to a higher position outside of the PP.
b. *Ik heb er ook op gerekend.
   I have there also on counted
   "I also counted on it."

c. *Ik heb ook op er gerekend.
   I have also on there counted
   intended: "I also counted on it."

However, we can exclude movement from the PP in a prepositional numeral based on the Phase-over-Phase Theorem (a Minimalist incarnation of the Complex NP Constraint): you cannot move out of a phase (here: the PP) that is immediately c-commanded by a phasal head (here: D°) (cf. 14) (Bošković, 2016). We can conclude that *er, although in principle able to licence unval[gender] on e, cannot occur in numeral PPs because it cannot move out of the numeral PP, leading to ungrammaticality of the sentence.

In Dutch, één “one” can inflect, and this is phonologically similar to adjectival inflection (schwa-inflection). One could therefore ask why, instead of inserting de in sequences like rond de één jaar “around one year”, één does not inflect. On deeper inspection, it appears that inflection on één is not inflection for [gender] but definiteness inflection (Barbiers, 2005). Consider the following examples (24).

(24) a. een goed boek     c. één boek     e. het één-e boek
    a nice book       one book       the one-INFL boek

b. een goed-e film     d. één film     f. de één-e film
    a nice-INFL movie one movie       the one-INFL movie


If één-inflection was dependent on the gender of the noun, the expectation would be that (24d) shows the same inflection as (24b). This is not the case; instead, één only inflects when it is combined with a definite determiner. Thus, één-inflection is not inflection for [gender] and therefore not able to license an empty noun by valuing its unval[gender]. It is still necessary to insert an element valued for [gender].

It appears that in some prepositional numerals leaving out the article leads to a less strong ungrammaticality effect. An example is given in (25a, cf. 25b).

(25) a. rond 20 mensen
      around 20 people
      “around 20 people”

b. * in 20 mensen
    in 20 people
    intended: “more than 20 people”
I tentatively suggest that this is because a preposition like rond is compatible with dates and times (in which the article is obligatorily absent) (e.g. rond 8 augustus lit. “around 8 August”) whereas a preposition like in is not (“in 8 augustus). This superficial similarity could increase the acceptability of the date/time-compatible prepositions. More research is necessary to test this suggestion, however.

There is one prepositional numeral construction that is slightly different from the others, namely tussen n en m “between n and m”. In Dutch, this expression is grammatical with both one or two articles (26).

   (26) a. Marie heeft tussen de 20 en de 30 boeken gelezen.
       Marie has between the 20 and the 30 books read
   b. Marie heeft tussen de 20 en 30 boeken gelezen.
       Marie has between the 20 and 30 books read

   “Marie read between 20 and 30 books.”

This can readily be explained by the theory if the coordination in these examples applies to different levels: in (26a), the DPs are coordinated, hence there are two structures in which unval[gender] needs to be valued, leading to insertion of de in both structures. In (26b), the NPs are coordinated, which means that de can value both unval[gender] features via Agreement. There is, however, a certain numeral PP structure that the current theory is unable to capture. An example is (27).

   (27) Marie is tussen de 6 weken en (de) 3 maanden zwanger.
       Marie is between the 6 weeks and (the) 3 months pregnant
   “Marie is between 6 weeks and 3 months pregnant.”

Assuming that the numeral expression is formed independent of the noun that it modifies, it should not be possible to have this noun within the numeral PP. At this point, I do not have a satisfactory explanation for this structure, and I will leave it for future work.

5. Conclusion
In this paper I have looked at prepositional numerals in Dutch and English. Following Corver & Zwarts (2006), I assumed that the numeral PP merges as the complement of the noun that it modifies, and I proposed that this structure includes an empty noun with an unvalued feature. Based on Barbiers (2005), I proposed that this feature needs to be licensed by valuing its unvalued feature via Reverse Agree (Wurmbrand, 2014). In English, this feature is unval[count]; in Dutch, it is unval[gender]. This results in the insertion of the gender marked article de in Dutch numeral PPs. The numeral PP then moves to Spec,NumP as an instance of DP-internal predicate displacement to derive the surface word order. Thus, cross-linguistic variation within prepositional numerals can be reduced to variation in the lexicon, in line with Chomsky (1995).

Future research might look into expanding the proposal outside of the domain of prepo-
sitional numerals. For instance, one could look into numerals modified by adverbials (e.g. ongeveer 10 “approximately 10”), that show different syntactic properties than prepositional numerals (obligatory absence of the article in Dutch; adverbials can occur to the left and to the right of the numeral). At first sight, an alternative analysis than the one proposed here seems necessary, but further investigation is needed to confirm this. Another option for further research is to take a bigger sample of languages and compare the internal properties of numeral PPs with that of noun ellipsis constructions. If there are parallels between these constructions in other languages as well, then this would be substantial support in favor of the analysis I suggested.

Received March 2017; accepted July 2017.

References