On predicate numerals

Norbert Corver
Utrecht University/Uil, OTS

1. Numerals as nominals

The true grammatical nature of what are traditionally called numerals is hard to grasp. Numerals like five, eight, twenty, etcetera denote a number. This number either designates an object, as in (1a), or a (quantitative) property of some object, as in (1b) (cf. Heeroma 1948). Given the fact that denotation of a property is a characteristic of adjectives, one sometimes classifies a numeral like tien in (1b) as an adjective or adjective-like element. Such an analysis would imply that the lexical item tien in (1a), which — in combination with the determiner — has referential force, is of a different categorical type (viz. N) than the lexical item tien in (1b), which denotes a quantitative property and consequently might be classified as an adjective.

(1) a. Ik heb een getal onder de tien in mijn hoofd
    I have a number below the ten in my head
b. Er staan tien koeien in de wei
    there stand ten cows in the meadow

While in traditional grammars the discussion about the classification of numerals centers around the question “Is it a noun, or is it more of an adjective?”, in generative grammar the discussion concerns the functional-lexical distinction. Although there will be little discussion about the categorial status of tien in (1a) — a lexical category of the class N —, unanimity in classification will be far less for the numeral tien in (1b). Jackendoff (1977), for example, claimed that a numeral like tien in (1b) is simply a noun, i.e. a lexical category. Selkirk (1977), on the other hand, argued that a numeral like tien in (1b) should be analyzed as a QP, i.e. a functional category. The two analyses are, in somewhat simplified form, represented in (2):

(2) a. [NP [NP tien] [N koeien]]
   b. [NP [QP tien] [N koeien]]
Most of the more recent analyses of the syntax of numerals have adopted the functional (i.e. QP) analysis (cf. e.g. Abney 1987; Barbiers 1990; Cardinaletti and Giusti 1991; Doetjes 1997). They differ, however, from Selkirk’s analysis in that they take Q to be a functional head that selects an NP as its complement:

(3) \[ [Q \{tien\} [NP koeien]] \]

It is quite obvious that an analysis according to which the (lexical) numeral \textit{tien} in \(1\text{(a)}\) is of a different categorial type than the homonymous functional one in \(1\text{(b)}\) leads to an increase of lexical items in our lexicon. It is more parsimonious if \textit{tien} in \(1\text{(a)}\) and \textit{tien} in \(1\text{(b)}\) are realizations of one and the same (mental) lexical item. Given the clear lexical (i.e. nominal) status of \textit{tien} in \(1\text{(b)}\) — consider, for example, its co-occurrence with the definite article \textit{de} — one comes to the conclusion that \textit{tien} in \(1\text{(b)}\) is a nominal element as well. If so, the difference in interpretation (say, referential \textit{versus} non-referential) between \textit{tien} in \(1\text{(a)}\) and \textit{tien} in \(1\text{(b)}\) should not be explained in terms of some categorial distinction but in terms of something else. Here I think the proposal by Stowell (1991) and Longobardi (1994) that nouns — like verbs, adjectives and some prepositions — function as predicates, provides a way of dealing with the interpretative contrast between \(1\text{(a)}\) and \(1\text{(b)}\). In both \(1\text{(a)}\) and \(1\text{(b)}\), the lexical item \textit{tien} is a (nominal) predicate. In \(1\text{(a)}\), the referential interpretation \textit{tien} results from its combining with the definite article (i.e. D) \textit{de}. In other words, it is the structural context (i.e. the DP-structure on top of the number-designating noun) which turns the nominal element into a referential argument of the prepositional predicate \textit{onder}.

Consider, next, the nominal element \textit{tien} in \(1\text{(b)}\). In line with the ideas of Stowell and Longobardi, I will assume that \textit{tien} simply is a predicative noun. This predicative, number-designating noun itself is not “closed off” by a reference-determining definite article. As a consequence, the numeral \textit{tien} in \(1\text{(b)}\) does not function as a referential argument of some other lexical predicate. It rather predicates the number-property “ten” over the nominal element \textit{koeien}. The complex phrase \textit{tien koeien} in \(1\text{(b)}\) functions as an argument of the verb \textit{staan}; this argumental status is plausibly due to the presence of an empty indefinite determiner on top of the phrase \textit{tien koeien}:

(4) \[ [DP e [ZP tien koeien]] \]

Thus far, I have made two claims on the basis of the examples in \(1\). First of all, both \textit{tien} in \(1\text{(a)}\) and \textit{tien} in \(1\text{(b)}\) are of the categorial type N(oun). Secondly, being a lexical category, it functions as a predicative element. The following question then arises: What is the internal syntax of the phrase \(ZP\) \textit{tien koeien} in \(4\)? That is, in what way are the two juxtaposed nominal elements in \(4\) combined? One analysis would be to say that the nominal element \textit{tien} Merges as a left branch element with
the nominal element *koeien*. Alternatively, one could assume that the order *tien koeien* is a derived order. Under such an analysis, the nominal element *koeien* (pure-)Merges with *tien*, and the order *tien koeien* results from displacement of the numeral to a position to the left of *koeien*. In this article, I will defend the latter analysis: Numerals, that is, are nominal predicates that undergo DP-internal predicate displacement. Taking the theoretical position that the predication relationship between the “subject” *koeien* and the number-designating predicate *tien* is structurally represented in terms of a Small Clause (SC), one is led to an analysis in which there is a SC projected within the DP. This yields the structure in (5). The surface order is derived by leftward movement of the numeral to a position to the left of the subject of the SC.

(5) \[ [DP \text{D} [SC koeien [Pred tien]]] \]

This analysis of DP-internal numerals in terms of predicate displacement is reminiscent of recent analyses by Kayne (1994) and Den Dikken (1995) of such constructions as (6):

(6) a. cet idiot de Jean (French)
    that idiot of Jean
b. that idiot of a doctor (English)

In this *N de/of N*-construction, the first nominal element (*idiot*) functions as a (qualitative) predicate, which predicates over the second nominal element (*Jean/doctor*). Kayne (1994:106) argues that *idiot* originates as a clause-internal predicate and raises (across the subject *Jean*) to the specifier position of a P-headed clause that is selected by the functional head D.

(7) cet \[D/PP [NP idiot de [IP Jean io \[e\]j…\]]\]

Den Dikken also assumes that predicate movement is involved in the derivation of the *N de/of N*-construction. He proposes the following derivation (see also Bennis, Corver & Den Dikken 1998 for discussion):

(8) \[ [DP that [FP idiot \[F (= of) + X_i (= a) [XP doctor [X_t t_j]]]]]] \]

In this analysis, predicate movement as found in (8) is taken to be an A-movement operation. What characterizes this movement operation is that the inverted nominal predicate skips an intermediate A-position, viz. that of the Small Clause subject (i.e. XP). Hence, the movement of the nominal predicate appears to be a non-local A-movement. As Den Dikken points out, however, the predicate movement is local if one adopts Chomsky’s (1993) locality theory in terms of equidistance. Under this theoretical proposal, the moved predicate can cross the subject as long as the two nominals are technically equally far away from the
Norbert Corver

predicate’s extraction site. Under Chomsky’s assumptions, this situation is obtained by the application of a domain-extending head movement operation that creates a minimal domain that contains both the raised predicate and the small clause subject. Den Dikken argues that in the case of DP-internal predicate inversion, the requisite domain extending head-movement operation consists of raising of the functional head (X) of the small clause to a higher functional head (labeled here as ‘F’). He further claims that the element de/of is a nominal copula, which surfaces as a result of X-to-F raising; in fact, this nominal copula is the (nominal) equivalent of the verbal copula to be, which obligatorily appears in predicate inversion structures in the clausal domain (e.g. I consider the best candidate *(to be) John); cf. Moro (1991).

Comparing the \(N \text{ de/of } N\)-constructions in (6) with the noun phrase tien koeien in (1b), we immediately observe that they are not similar: they differ in the appearance of a linking, preposition-like element. The former construction displays such an element, the latter does not. Given this, one might jump to the conclusion that a construction like tien koeien does not feature predicate movement after all. However, one should not jump to conclusions to hastily on the basis of superficial asymmetries. If there is a language where the parallelism between \(N \text{ de/of } N\)-construction and the Num + N-construction is complete, one should try to uphold a similar syntactic treatment of the two construction types also for those languages where parallelism is superficially not complete. In fact, such an approach to the study of nominal construction types is in line with what Chomsky (2001) calls the Uniformity Principle, which states the following: “In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.”

In the next section, I will consider numeral constructions in Romanian. As we will see, this language displays the pattern NUM (= number-denoting Noun) de N, i.e. a pattern which is superficially very similar to the \(N \text{ de } N\)-pattern in (6).

2. DP-internal Predicate displacement of numerals

Consider the following examples:

\[(9) \quad \begin{array}{l}
\text{a. treizeci de lei} \\
\text{thirty of lei} \\
\text{‘thirty lei’} \\
\text{b. douăze ci și unu (de kilograme) de mere} \\
\text{twenty and one (of kilos) of apples} \\
\text{‘twenty one apples’}
\end{array}\]
This pattern is, clearly, very similar to the \(N \text{ de/of } N\)-pattern. In (6a), it is a quality-designating nominal predicate that is linked to the subject by the element \(de\). In the examples in (9), this linking element relates a number-designating predicate nominal to the subject of the predication relationship.\(^1\) The parallelism between the two construction types is suggestive, of course, for a similar treatment. This implies that an example like (9a) has the following derived structure:

(10) \[
\left[ FP \text{ treizeci} \left[ F \text{ (= de)+} X_i \left[ XP \text{ lei } X' \text{ ti tj} \right] \right] \right] \]

The NP \text{ treizeci} starts out in the predicate position of the DP-internal Small Clause (XP) and raises to \([\text{Spec,FP}]\) after the domain-extending head-movement operation \(X\text{-to-F}\) has applied.

Adopting Chomsky’s Uniformity Principle, one is led to an analysis in which the noun phrase \text{ tien koeien} in (1b) also features a predication relationship and a predicate displacement operation which places the numeral predicate in a position to the left of the subject. And, of course, if this analysis applies to the Dutch sequence \text{ tien koeien} in (1b), then it plausibly extends to the English equivalent \text{ ten cows}. The only difference between Dutch and English, on the one hand, and Romanian, on the other hand, seems to be the presence of the linking element \(de\).\(^2\)

In the Romanian examples in (9), \(de\) is obligatorily present: \text{ treizeci *(de) lei}. In Dutch and English, on the contrary, the linking element is obligatorily absent: \text{ tien (*van) koeien} (Dutch); \text{ ten (*of) cows}.

As a matter of fact, the situation is even more complicated: even though the English and Dutch \text{ NUM+N}\-constructions are very much alike superficially, certain empirical facts suggest that at a more abstract level they are different. A striking asymmetry, for example, is the one between (11) and (12):

(11) a. Jan heeft [[tien of zo mensen] uitgenodigd
    b. Jan heeft [tien mensen of zo] uitgenodigd

(12) a. John solved the problem in [[twenty or so] minutes]
    b. She visited [[fifty or so] theatres]

Consider, first, the English examples in (12). The numeral is coordinated with the element \(so\) by means of the coordinating conjunction \(or\). The interpretation of the complex \text{NUM or so} is that of approximation. The string \text{ twenty or so} receives the interpretation: “twenty or in the vicinity of twenty”. This approximative reading arguably results from the presence of a disjunctive coordinator in combination with
the indefinite element *so*. The disjunctive coordinator expresses an alternative, adversative relation: what is expressed by the left conjunct is considered to be an alternative for what is expressed by the right conjunct. Being coordinated with a numeral (i.e. quantity-denoting noun), the indefinite element *so* receives a quantitative interpretation, under the assumption that members that stand in a coordinate relationship are of the same semantic type. Thus, *so* in a way inherits its quantitative meaning from the left member of the coordinate structure. Since *so* is indefinite — i.e. it does not designate a definite number — we get the approximative reading “twenty or a number like twenty” (cf. (12a)).

As shown by the following examples from Dutch, this approximative reading of the sequence of *zo* (‘or so’) can also apply to a non-quantitative property:

(13) a. [Bakker of zo] wordt Jan later  
   Baker or so becomes Jan later  
   b. [Echt slim of zo] is hij niet  
   Really smart or so is he not

In (13a), *zo* receives the meaning: “something close to the qualitative property of being a baker”, and in (13b) *zo* has the meaning: “some quality which is close to being smart.” Although a full analysis of the semantics of approximation falls beyond the scope of this article, it is clear that the sort of approximative reading (e.g. quantitative, qualitative) we get in these examples is dependent on the meaning properties of the left member.

Observe that in the examples in (13), *zo* is coordinated with an expression (*bakker, echt slim*) that has a predicative function. This is very clear, of course, when the sequence of *zo* is omitted:

(14) a. Bakker wordt Jan later  
   b. Echt slim is hij niet

Under the assumption that the members of a coordination must be of the same semantic type (i.e. argument & argument; predicate & predicate), we come to the conclusion that (approximative) *zo* is predicative as well. The approximative reading is the result of the disjunctive coordination configuration in which this element occurs. In structural contexts where this coordinator is absent, the predicative element *zo* does not get the approximative reading. Compare e.g. the following examples with those in (15):

(15) a. Zo wordt Jan later  

   so becomes Jan later  

   ‘Jan will become like that’
On predicate numerals

b. Zo is hij niet
   so is he not
   ‘He is not like that’

English so, just like its Dutch equivalent zo, has a predicative function, as is exemplified in (16):

(16) a. John is *very busy* and he expects to be so for the next four hours
    b. John became a linguist and remained so for the rest of his life

In (12), the pro-predicate so is coordinated with the numeral by means of the disjunctive coordinator or. Just like in Dutch, it is this disjunctive coordination configuration that yields the approximative reading.

Having determined that so (and Dutch zo) is a pro-predicate, let me further add that this element typically substitutes for a phrasal constituent (YP) and not for a head (Y). Compare, for example, the following examples from English:

(17) a. John was fond of rats and remained so for the rest of his life
    b. *John was no longer fond of rats but he remained so of mice

Given the X-max-status of so, it can be concluded that the numeral twenty/fifty in (12) is a phrasal constituent (i.e. YP) as well; this under the generally held assumption that conjuncts of a coordinate structure must be of the same projection level; that is, a phrasal constituent can be coordinated with another phrasal one (i.e. XP & YP) but not with a head-like constituent (i.e. *X & YP).

Now that we know that a string like twenty or so involves a conjunction of two predicative XPs, we come to the following analysis of a noun phrase like twenty or so minutes in (12a).

(18) [FP [twenty or so]j [F+F+Xi [XP minutes [X+ti tj]]]]

The phrase twenty or so is a maximal phrase (arguably a Conjunction Phrase) containing two maximal phrases, twenty and so. Just like in the Romanian pattern in (10), the (complex) numeral has undergone XP-movement to [Spec,FP], after X has raised to F. As opposed to (10), the English example does not feature an overt copular element.

Let us now turn to the contrast between the Dutch examples in (11) and the English examples in (12). The ill-formedness of (11a) suggests that a structural analysis like (18) is excluded for Dutch. But what then is the correct analysis for the Dutch sequence NUM+N? For my answer to the question as to what underlies this contrast I will first consider the syntax of pseudopartitive constructions (cf. Selkirk 1977).
3. Towards a typology of numeral constructions

In Corver (1998), I argued that pseudopartitive constructions like English *a bottle of wine* and French *une bouteille de vin* involve the phenomenon of predicate displacement: the quantity designating nominal predicate *bottle/bouteille* starts out as a DP-internal Small Clause predicate that takes the nominal *wine/vin* as its subject (i.e. external argument). The quantitative predicate is moved leftward across the subject, and the copular element *of/de* surfaces as a result of the process of predicate inversion. Schematically (for English):

(19) \[DP \text{ a } [FP \text{ bottle}] [F (~= of) + X_1 [XP \text{ wine} [X_2 \text{ t}_1 \text{ t}_2]]]]

The predicative status of *bottle* is corroborated by the existence of approximative constructions like *a [bottle or so] of wine*, where the pro-predicate *so* is conjoined with the quantity designating NP *bottle*.

What is interesting is that in a language like Modern-Hebrew there are two other patterns of the pseudopartitive construction besides the one featuring the nominal copula *shel* (e.g. *bakbuk shel yáyin*; bottle of wine). These two patterns are: (i) the apposition pattern (cf. (20)) and (ii) the Construct State pattern (cf. (21)); see Glinert (1989). Both patterns are characterized by the absence of a linking copular element; the two nominals are, so to speak, juxtaposed.

(20) Apposition pattern  (21) Construct State pattern

a. bakbuk yáyin  a. bakbuk yáyin
  *bottle wine*   *bottle wine*
  ‘a bottle of wine’ ‘a bottle of wine’

b. tipa mayim  b. tipat mayim
  *drop water*   *drop water*
  ‘a drop of water’ ‘a drop of water’

Although the patterns (20a) and (21a) are superficially alike, the morphological contrast between (20b) and (21b) shows that the two types of (juxtaposed) pseudopartitive constructions should be distinguished. In (20b), the free (i.e. non-construct) ending *-a* (feminine, singular) is attached to the stem *tip*, whereas in (21b) the construct ending *-at* (feminine, singular) is attached to it.

Another contrast between the Apposition pattern and the Construct State pattern concerns the syntactic distribution of expressions of approximation of the type *or + numeral*. As shown in (23), the expression of approximation must follow N2 in the Construct State variant of the pseudopartitive construction, whereas in the apposition variant it is preferred to have it in between N1 and N2 (cf. (22)).
On predicate numerals

(22) a. *tipa o shtayim mayim  
   b. tipa o mayim shtayim  
   drop or two-FEM water  
   'a drop or two of water'

(23) a. *tipat o shtayim mayim  
   b. tipat mayim o shtayim  

It does not seem implausible to analyze the numeral (shtayim) following the coordinating conjunction o, as an incomplete (i.e. ellipted) phrase; that is, the sequence tipa o shtayim stands for “(a) drop or two drops”. Being an (ellipted) phrasal constituent, it must be coordinated with a constituent of the same hierarchical level given the standard assumption that conjuncts must have the same projection level. In other words, the left conjunct tipa in (22) is phrasal as well. The (derived) structure of (22a) is then as follows:

(24) [DP D[FP [ConjP [tipa] o [shtayim pro] F+X [XP mayim [ti tj]]]]]

In fact, this is the pattern that is also found with the shel-variant of the pseudopartitive construction: e.g. bakbuk o shnayim shel mayim (bottle or two of wine).

The ill-formedness of (23) follows from the traditional generalization that the raised noun (N) in Construct State constructions must always be string-adjacent to the second nominal. It is impossible to have phrasal material in between the raised noun tipat and the nominal mayim. As illustrated in (23b), the expression of approximation remains stranded in a position following mayim. The only element moved out of (the specifier position of) the Conjunction Phrase is the N tipat. I assume that this raised Noun adjoins to the F+X-complex that is the result of X to F-raising. The structure of the (23a) is then, schematically, as follows:

(25) [DP D[FP [tipat]+[F+X [XP mayim [X ti [NP shtayim pro]]]]]]

Summarizing, in Modern-Hebrew there are three patterns of the pseudopartitive construction: (i) a shel-pattern, (ii) an apposition pattern and (iii) a Construct State pattern. The first two patterns involve movement of a phrasal nominal predicate (i.e. NP). The Construct State pattern, on the contrary, is characterized by the application of N-raising to the nominal predicate.

On the basis of the distribution of expressions of approximation, it was concluded in Corver (1998) that Dutch pseudopartitive patterns (e.g. een flesje wijn; a bottle wine, ‘a bottle of wine’) are derived by predicate movement of the N-raising type. Just like in the Modern-Hebrew example (23), the approximative phrase can never move along with the first nominal element of the coordinate structure:

(26) a. *Na [een flesje of twee wijn] voel ik me altijd veel beter
    After [a bottle or two wine feel I myself always much better
    b. Na [een flesje wijn of twee] voel ik me altijd veel beter
The structure of a Dutch pseudopartitive construction like *een fles wijn* will then be the following:

\[
(27) \quad [DP \: \text{een} \: [FP \: \{\text{fles} \: + \: \{F + \text{X}\} \} \: [XP \: \{\text{wijn} \: \{\text{X}\} \}]])
\]

We can now return to the numeral constructions and address the question of what “typological” variation is found for this construction type. What is interesting is that we find the same patterns as those found for the pseudopartitive construction. That is:

\[
(28) \quad \begin{align*}
\text{a. treizeci de lei} & \quad \text{(NUM de N-pattern; Romanian)} \\
& \quad \text{thirty of lei} \\
& \quad \text{‘thirty lei’} \\
\text{b. shisha rofim} & \quad \text{(NUM (=NP) N-pattern; Modern-Hebrew)} \\
& \quad \text{six (free form) daughters} \\
\text{c. shney dagim} & \quad \text{(NUM (=N) N-pattern; Modern-Hebrew)} \\
& \quad \text{two (CS-form) fish}
\end{align*}
\]

For the derivation of the pattern in (28a), see (10) in Section 2. On analogy with the apposition pattern in (24), I will assume that pattern (28b) is derived by application of predicate displacement of the XP-movement type.

\[
(29) \quad [FP \{\text{shisha} \} \: [F + \text{X} \: [XP \: \{\text{rofim} \: \{\text{X}\} \}]])
\]

As suggested by the well-formedness of the examples in (12) — recall that coordination with the pro-XP so suggests phrasal status of the numeral — the English NUM+N-pattern is of the same type as (28b); whence (18).

Having our typology of numeral construction types, we can finally interpret the Dutch pattern *tien koeien* (ten cows). The distribution of the approximative expression of *zo* in (11) is very suggestive. This expression can never move along with the numeral to a position preceding the second nominal (N2). The sequence of *zo* must be stranded, just like in the pseudopartitive patterns featuring N-raising (i.e. the Construct State pattern). From this I conclude that the Dutch construction *tien koeien* has the following derived representation:

\[
(30) \quad [DP \: D \: [FP \{\text{tien} \} \: [F + \text{X} \: [XP \: \{\text{koeien} \: \{\text{X}\} \}]])
\]

Just like in the Construct-State variant of the pseudopartitive construction, N-raising has applied to the predicate nominal. The raised Noun gets adjoined to the complex head [F+X].
4. Conclusion

In this article, I have argued that: (a) DP-internal numerals are predicate nominals that undergo predicate displacement; (b) Predicate displacement can be of the XP-movement type or X-movement type; (c) pseudopartitive and numeral constructions share the typology of realization patterns.

Notes

1. The presence of an underlying *de* in nominal constructions featuring a quantifying element (e.g. a numeral like *dix* (‘ten’) or a quantifier like *beaucoup* (‘many’)) was already argued for in the seventies by Jean-Claude Milner for French (see e.g. Milner 1978). The linking element *de* is obligatorily present in a string like *beaucoup de livres* (many of books; ‘many books’) but obligatorily absent if the quantity-designating element is a numeral: *dix *(‘de’) livres (ten of books; ‘ten books’). Presence of an underlying *de* in the latter string is suggested, however, by *en*-extraction, as in: *J’en ai lu beaucoup* (I there-of read many) and *J’en ai lu dix* (I there-of read ten). The clitic *en* only substitutes for a nominal element introduced by *de*.

2. Not all numerals are linked to the following noun by means of the element *de*. It only appears with numerals from 20 onwards. The string *four apples*, for example, has the following translation in Romanian: *patru mere* (four apple*pl*).

References

Dikken, M. den (1995) Copulas, lezing gepresenteerd op GLOW 18, Tromsø, Ms., Vrije Universiteit Amsterdam/HIL.