Split DP Topicalization and the role of interfaces

Ion Giurgea
Université Paris VII
igiurgea@linguist.jussieu.fr, giurgeaion@yahoo.com

Abstract.
Split DP Topicalization, found in various languages such as German, Romanian, Albanian or Hungarian, looks as a DP with the determiner (usually an indefinite or quantifier) stranded in the base position and its nominal complement fronted to a sentence-initial position, with topic interpretation. This structure is problematic because it seems both to require and to exclude an analysis by movement. Based on evidence from Romanian and German, I show that the movement analysis is correct, and that all the properties which seem to plead against it result from PF processes. The theoretical implications of this analysis are that post-syntactic morphology is necessary, and that certain agreement phenomena, which may be described as “PF-agreement”, take place late, inside a phase domain, by the time the phase is completed and the phase domain becomes opaque.

1. The puzzle
1.0 Introduction
I discuss examples of the form:

(1) a. Cărți am cumpărat multe
    books I-have bought many
b. Bücher habe ich viele gekauft
    books have I many bought

This type of configuration, called “inverted split DP”¹ in Fanselow’s typology of split-DPs (in contrast to structures of the type “many have I bought books”, which are called “simple splits”), has been a long standing problem in the generative literature, because it seems both to require and to exclude a movement analysis.

1.1. Arguments against movement
(a) In some cases, the topical N(um)P and the D put together cannot form a constituent:
- The determiner left in the base position has special “pronominal” or pre-[Nc] forms, impossible before an overt noun:

(2) a. Am cumpărat un disc
    I-have bought a CD
b. Am cumpārat unul
    I-have bought one
c. Discuri am cumpārat unul
    CDs have I bought one
d. *Am cumpārat unul disc(uri)
    have I bought one CD(s)

(3) a. Ich habe kein Geld.
    German

¹ This construction is found in many languages, but I will concentrate on Romanian and German.
I have no money
b. Ich habe keines
I have none
c. Geld habe Ich keines
money have I none
d. *Ich habe keines Geld
I have none money

- Number mismatch is possible between the determiner and the topical NP:

(2)  
  a. Kinder hat er nur eines   
  children has he only one
  b. *Er hat ein(es) Kinder   
  he has one children

- The topical NP can have a determiner of its own: the indefinite article *ein*, in some varieties of German:

(5)  
  a. Ein Auto hat er keines  
  a car has he none
  b. *Er hat keines ein Auto 
  he has none a car

- In German, adjectives in the topical N(um)P don’t have the “weak” inflection, used when preceded by a determiner, but the “strong” inflection, used in the absence of a determiner:

(6)  
  a. Er hat keine polnischen Gänse 
  he has no Polish-*en* geese
  b. Polnische Gänse hat er keine 
  Polish-*e* geese has he no
  c. *Er hat keine polnische Gänse   
  he has no Polish-*e* geese
  d. Er hat polnische Gänse        
  he has Polish-*e* geese

- In Romanian, some cardinals obligatorily insert *de* “of” before their nominal complement. In the “split” construction, *de* is absent:

(7)  
  a. Au venit douăzeci de copii 
  have come twenty of children
  b. (*De) copii au venit douăzeci (*de) 
  (of) children have come twenty (of)
c. *Au venit douăzeci copii
have come twenty children

- In some varieties of German (or at least for some speakers), prepositions may appear twice, before each element:

(8) a. In Schlößern habe ich in keinen gewohnt
in castles have I in none lived
b. *Ich habe gewohnt in keinen in Schlößern
I have lived in no in castles

- In German the indefinite welche cannot appear followed by an overt noun, but may appear in the split construction:

(9) a. Unbeschädigte Exemplare habe ich kaum noch welche (Van Riemsdijk 1989)
undamaged copies have I hardly still some
b. Hast du *(irgend)welche unbeschädigten Exemplare?
have you any undamaged copies
have you some
c. Hast du welche?

(b) The elements allowed in the topical and in the base position may occur in these positions in the absence of an “associate” (i.e. outside “split-DP”):

(2)´ a. Discuri am cumpărat unul
CDs I-have bought one
b. Am cumpărăt unul
I-have bought one
c. Discuri am cumpărăt
CDs I-have bought

(3)´ a. Geld habe Ich keines
money have I none
b. Ich habe keines
I have none
c. Geld habe Ich
money have I

(7)´ a. Copii au venit douăzeci
children have come twenty
b. Au venit douăzeci
have come twenty
c. Copii au venit
children have come

From the formulation in (b) it follows that phrases which cannot occur independently in these positions should not be able to occur in the split-construction either.
For the topic position, this may seem to be true in Romanian and some varieties of German:
(10) a. *Fată a venit / *A venit fată
girl has come girl
b. A venit o fată
has come a girl
c. *Fată a venit una
girl has come one

Some varieties of German contradict this consequence of the formulation in (b): in these dialects, one can find in the topic position forms that require a split:

(11) Guten Professor kennt sie ?? (einen/keinen)   Southern German
good teacher knows she (one/none)

For the base position, there is positive evidence in Romanian: if a determiner cannot combine with an empty N, it cannot appear in the split construction:

(12) a. Au venit nişte *(fete)     Romanian
have come some(s’m) girls
b. *Au venit nişte
have come some(s’m)
c. *Fete au venit nişte
girls have come some(s’m)

1.2. Arguments for movement
(a) The strongest evidence for movement comes from locality: the relation between the topical NP and the “associate” shows island violation effects typical for movement. Thus, in Romanian the “associate” must be an object or a postverbal subject:

(13) a. Doctori am consultat mulţi
physicians I-have consulted many
b. * (??) Doctori am vorbit cu mulţi
physicians I-have spoken with many
c. Candidaţi au telefonat mulţi
candidates have called(on the phone) many
d. * Candidaţi mulţi au venit (impossible with the intonation characteristic of topics)
candidates many have called

This is expected under a movement analysis since PPs and preverbal subjects are islands for extraction in Romanian.
We should however note that the locality conditions aren’t always the same as for the wh-movement: *wh- extraction out of DP is generally bad, even from object and postverbal subject positions, except for possessor extraction:

(14) a. *De ce ai consultat doctori? (in the interpretation “what sort of physicians…”)
of what you-have consulted physicians
b. * Despre cine ai cumpărat o carte?
   about whom you-have bought a book

c. (?) Cui i-ai văzut cartea în librărie?
   whom.D DCl-you-have seen the-book in (the) bookstore

The locality conditions don’t match those of the extraction of a whole DP either:

(15) a. * (?) Copii/La copii (le-)am dat cadouri multora
   children/D children (DCl-)I-have given presents many.D
   b. Cui i-ai dat cadouri ?
   whom.D DCl-you-have given presents

Both in Romanian and German the topicalization obeys clausal islands (wh-, adjunct- and
“complex NP” islands):

(16) a. *Unbeschädigte Exemplare wollte er wissen wer noch zwei auf Vorrat hat     German
   undamaged copies wanted he to-know who still two on stock has
   (Van Riemsdijk 1989)
   b. *Calculatoare se supără dacă nu luăm trei
   computers Refl gets-angry if not we-take three
   Romanian
   c. ?(%) Studenti nu cunosc nici un profesor care să aibă mai puțin de zece
   students not I-know no professor which Subj has less than ten
   d. ?(?) Poezii se supără dacă nu învăț măcar trei
   poems Refl gets-angry if not I-learn at-least three
   e. *Studenten kenne ich keinen Professor der weniger als zehn habe        German
   students know I no professor which fewer than ten has.Subj

In German, the constraints are similar to those of wh- movement:

(17) a. ?? Ein Experte hat keiner den Schaden geprüft     (Van Riemsdijk 1989)
   an expert has no-one the damage proven
   b. * Was hat für ein Mann den Schaden geprüft?
   what has for a man the damage proven “What sort of man has established the damage”
   (Fanselow and Cavar 2002)

(18) * Mädchen hat er mit vielen getanzt   (Van Riemsdijk 1989)
   girls has he with many danced

However, at least for some speakers, there are cases where splits possible while other kinds of
subextraction are forbidden. This is the case of genitive and dative objects:

(19) a. Interessanten Büchern über Polen ist hier noch keinen ein Preis verliehen worden
   interesting books-D about Poland is here yet none-D a price awarded been
   b. *Über Polen ist hier noch keinen Büchern ein Preis verliehen worden
   about Poland is here yet no-D books-D a price awarded been
   (Fanselow and Cavar 2002)
(b) A second indication that the topic and the associate are syntactically, not just semantically related is the impossibility of a lexical N in the base position, even when the semantics or intended meaning would allow one:

\[(20)\]
\[\begin{align*}
    a. & \text{ Fete } \text{au venit } \text{Maria şi Cristina} & \text{Romanian} \\
    & \text{girls have come M. and C.} \\
    b. & \text{Animale îi plac numai câinii} & \text{German} \\
    & \text{animals him.D like only dogs} \\
    c. & \text{Raubvögel kennt er nur Bussarde} & \text{Fanselow 1988} \\
    & \text{birds of prey knows he only buzzards}
\end{align*}\]

With hanging topics with an introductory element (ca in Romanian, corresponding to the English \textit{as for}), there is no such constraint:

\[(21)\]
\[\text{Ca fete, au venit numai Maria şi Cristina} \quad \text{“As for girls, only Maria and Cristina came”} \]

(c) The literature about split DPs often invoked reconstruction facts as an argument for movement:

\[(22)\]
\[\begin{align*}
    a. & \text{Bücher über einander hatten die Männer keine geschrieben} & \text{German} \\
    & \text{books about each-other had the men no written} \\
    b. & \text{Poezii despre sine, (el) a scris zece} & \text{Romanian} \\
    & \text{poems about himself (he) has written ten}
\end{align*}\]

\[(23)\]
\[\begin{align*}
    a. & \text{Articole despre președinte, au adunat consilierii lui, destule} & \text{Romanian} \\
    & \text{articles about (the) president have gathered his counselors enough} \\
    b. & \text{*Articole despre președinte, a adunat el, destule} & \text{Romanian} \\
    & \text{articles about (the) president has gathered he enough}
\end{align*}\]

However, this argument is not so strong, since reconstruction is also found with base generated topics, such as the \textit{as-for} topics and the English Left Dislocation construction (see (24)), which is, according to Frey (2005), an instance of base generation. Indeed, (25) shows that this structure is insensitive to islands and allows overt (lexical) Ns in the base position, in the appropriate semantic relation (part-whole):

\[(24)\]
\[\begin{align*}
    a. & \text{As for books about each other, I don’t think they ever wrote any.} \\
    b. & \text{Opinions of ? him/ himself, John has collected eleven so far} \\
    c. & \text{Opinions of John, he, has collected eleven so far.}
\end{align*}\]

\[(25)\]
\[\begin{align*}
    a. & \text{PhD students, I know a professor who has eleven} \\
    b. & \text{Sculptures by Michelangelo, I’ve seen Pieta and Moses}
\end{align*}\]

(d) We find further evidence for movement in (some varieties of) German:

- Topical NPs occupy the so-called SpecCP position of V2 clauses, and not a left dislocation or hanging topic position. This position is usually occupied by moved elements:
(26)  a. Bücher hat er viele gelesen  
books has he many read  
b. *Bücher er hat viele gelesen  
books he has many read  
c. * Maria glaubt daß Bücher er viele gelesen hat  
M. believes that books he many read has  

Even the left dislocation position (see (26)a), although exterior to the CP, appears to be created by movement (see Frey 2005), as the English “Topicalization”: it shows case agreement with the associate (which is a resumptive pronoun), and reconstruction effects. Hanging topics are distinguished from left dislocated elements by lack of agreement, lack of reconstruction and a different, comma-like intonation:

(27)  a. Den Hans → den mag jeder  
the.A H. him.A likes everyone  
(the arrow indicates “progresident intonation”)  
b. Hans, jeder mag ihn  
H. everyone likes him  

So, for a base generated topic we would expect the structure (26)b, the Hanging Topic.

- In some varieties of German we find number agreement between the topic and the quantifier in the base position (see (28)), and also case agreement (see (29)):

(28)  a. (Ein) Auto hat er nur eines  
(a) car has he only one  
b. * Autos hat er nur eines  
cars has he only one  
c. Autos hat er drei  
cars has he three  
d. * (*ein) Auto hat er drei  
(a) car has he three  

(29)  a. Alten Professor kennt sie schon einen  
old.Ac professor knows she already one.A  
b. Männern habe ich vielen geschrieben  
men.D have I many.D written  
c. Schrecklicher Morde ist er vieler beschuldigt worden  
horrible.G murders is he many.G accused been  
“He has been accused of many horrible murders” (Fanselow and Cavar 2002)  

- As Van Riemsdijk (1989) shows, if adjectives are found both in the topic position and in the base position, the adjective in the topic position must be closer to the N in the hierarchical order of adjectives (it must follow the other adjective in the usual A-N order when the DP is not split). The existence of such a constraint cannot be explained unless the adjectives were first merged inside the same DP:
(30) a. Ein Amerikanisches Auto kann ich mir kein neues leisten
   an American car can I me.D no new afford
b. * Ein neues Auto kann ich mir kein Amerikanisches leisten
   a new car can I me.D no American afford
Cf. c. ein neues Amerikanisches Auto
   a new American car
d. *ein Amerikanisches neues Auto
   an American new car

- The existence in some dialects of phrases (bare count singulars) in the top position which require an associate: see (11) above.

2. The problem in the context of minimalism
The arguments above prove the existence of a dependency in this structure.
In the current minimalist framework, with strict cyclicity and no SS-level, we cannot describe this dependency using the notion of a “representational chain” (as proposed by Cinque in a similar case, CLLD - see Cinque 1983).
However long distance dependencies can be described using the operation Agree. Adger and Ramchand (2005) have shown that syntactic dependencies involved in relativization may be described using Agree. They point out that the indication of movement is not merely the dependency itself, but dependency and identity.
But arguably we do find identity in interpretive features between the two “DP-splits”. We saw that part relations are not allowed (ex. (20)). The mismatches listed in section 1.1 concern uninterpretable features such as adjectival agreement, forms of determiners in the context of empty N, and number, which is not interpreted in the case of quantifiers, but its form in the base position is determined by the quantifier, while in the topic position it is either the same as in the base position (Agree) or a default (bare plural) (but see 3.2)
If only non-interpreted, purely formal features diverge, then the simplest assumption is that the divergences arise in the morphological component, and in so far as (narrow) syntax is concerned, the two elements (the topical NP and the null NP in the base position) are identical. This allows us to conclude that they are related by movement.
Moreover, Agree involves only grammatical features. What features could be responsible in this case for the dependency?
Adger and Ramchand (2005), for the Scottish Gaelic relativization, propose the features Λ (on the complementizer) and unvalued ID on the gap=pro, where ID, “identification”, may be either Λ, for bound variables, or ϕ. This replaces an analysis by null operator movement.

(31) C: Λ, uID:Λ ... pro: ID:- →
C: Λ, uID:Λ ... pro: ID:Λ =
λx ... x

But in our case the gap denotes a property. So we would be forced to introduce lambda abstraction over properties.
More importantly, the reconstruction effects cannot be derived in such an analysis. (We cannot appeal to the ellipsis-resolution mechanism if the gap is not an elided NP, but a property-type bound variable).
The same arguments argue against a Base Generation + Null Operator Movement analysis as proposed for CLLD by Iatridou (1995).
So I will adopt a movement analysis and try to give a morphological account for the facts illustrated in 1.1.

3. A morphological explanation for the facts illustrated in 1.1.

3.1. Full DP in the base position: pre-[Ne] forms of the determiners

The pre-[Ne] forms of the determiners (see (2)-(3)), the ban on the determiners which cannot combine with an empty N (see (12)) and the presence of determiners specialized to pre-[Ne] contexts (ex. (9)) may be derived as a consequence of the assumption that the morphological component treats base-generated empty Ns and traces of movement in the same way.

Indeed, according to the copy theory of movement, traces are not sui-generis objects, but phrases deleted at PF. Empty Ns are pro-Ns with a null phonological matrix (for an overt pro-N, see engl. *one*). So after “chain reduction”\(^2\) (copy deletion) applies, the two are the same object at PF: \([\text{Ne}]\).\(^3\)

3.2. Split DP and number: Number mismatch

Number mismatch only arises when there is a (count) singular in the base position. In this case, we find three constructions in the languages discussed here:

1. plural… singular (Romanian + one variety of German) (see ex. 2 and 4)
2. singular indef. article+NP … singular (a second variety of German) (see ex. 5)
3. bare singular NP … singular (a third variety of German) (ex. 11 and 26)

Construction 3 is unproblematic. Moreover, as we have seen (ex. (11) and (26)), it supports the movement hypothesis, since in the absence of an associate a bare singular is impossible in this position.

For 2 we may assume that topics must be at least NumP in these dialects, and singular Num is realized as the indefinite article in the absence of a preceding determiner.

For various reasons (see Munn and Schmitt 2001, 2002, 2004, Dobrovie-Sorin, Bleam & Espinal (2005))\(^4\) it has been suggested that the difference in distribution between bare count singulars and bare plurals reflects a different categorial status: bare plurals are NumPs, bare count singulars are NPs, and the singular counterpart of plural Num is the indefinite article.

Then the derivation will look as follows:

\[
(32) \quad \text{No movement: } [[D \text{ kein}] [\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow \text{kein Auto}
\]

\[
\text{Movement: } [[D \text{ kein}] [\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow
\]

\[
[\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow [[D \text{ kein}] [\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow
\]

\[
[\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow [[D \text{ kein}] [\text{NumP} [\text{Num} + \text{sg}] [\text{NP, Auto}]] \rightarrow
\]

\[
\text{ein Auto } \ldots \text{ keines}
\]

\(^2\) In Bobaljik’s (2002) PF-algorithm, chain-reduction precedes Vocabulary Insertion (it is the second step after assignment of precedence relations to syntactic nodes)

\(^3\) Neeleman and Weerman (1999) argue that in accordance with minimalist assumptions, we expect licensing of empty categories to be a PF phenomenon

\(^4\) Another argument for the analysis of bare plurals and singulars with the indefinite article as NumPs is their similar behaviour in predicative copular sentences:

(i) Fido and Guffy are dogs
(ii) Fido is a dog
(iii) *Fido is dog
To explain construction 1, number mismatch (plural .. singular), we start from the observation that in the dialects/constructions with number mismatch the determiners which allow the split construction are precisely those for which number appears to be uninterpreted: the Det is a quantity word which selects a particular number (cardinal or indefinite quantity word, including the negative in Romanian), or an indefinite for which the difference in number doesn’t usually add something to the interpretation (the negative in German); universal quantifiers, quodlibetics, alternatives and definite Ds are degraded:

(33) a. ?? Romane de Borges (le)-am citit pe acestea/fiecare/toate/altele Romanian novels by B. (AcCl)-I-have read Ac those/each/all/others
b. * Cărți ia-o / ia-le pe oricare books take-AcClsg. / take-AcClpl. any

Secondly, it is important to notice that plural also appears in base generation constructions, as we see in the English Left Dislocation, which, as shown in (25), is a base-generated (hanging) topic, and in as for topics in English and Romanian:

(34) Books/*Book about Paris, I have only one.
(35) a. As for prizes, he has only yet received one.
   b. Ca premii, până acum a primit doar unul Romanian as prizes until now has received only one

Based on these facts, I propose the following explanation for number mismatch:
(i) Num in structures with quantitative determiners has an unmarked value, interpreted as “(indefinite) quantity” or “sum of individuals” (see Dobrovie-Sorin 2006), or perhaps “divisive” (as proposed by Borer 2005). For count nouns, this unmarked value is the plural.
(ii) What is topicalized in Split-DPs is a NumP with an unmarked, +divisive value for Num. It is not merely the class denoted by the noun, but something that receives a quantity specification5. The quantity specification is part of the new information, always sitting in the base position.
(iii) In structures with quantitative determiners, Num (morphologically) agrees with the determiner.
(iv) The extracted Num escapes agreement with the Q because this agreement is late (hence the term “morphological”): it is realized on the phase domain by the completion of the phase, after all movements outside the phase have taken place.

We assume that the topicalized constituent is a NumP because if it were an NP, we would expect to find singular morphology, since singular morphology is simpler than plural (the singular form is often the bare root). We also have evidence for the assumption that the reason why the topic contains the Num level is not purely formal, but semantic: it is not the case that the topic position

5 Note that mass nouns are acceptable in this construction (at least in Romanian), appearing, of course, in the singular:
(i) Lapte am luat două sticle milk I-have taken(bought) two bottles
requires that the Num level should be realized, since bare count singulars, which we assumed to lack the Num level, may be topicalized\textsuperscript{6}, yet they cannot appear in the split construction:

\begin{enumerate}
\item a. Are maşină
\hspace{1cm} he-has car
\item b. Maşină are
\hspace{1cm} car he-has
\item c. Maşini/??maşină are doar una
\hspace{1cm} cars/car he-has only one
\item d. Maşini/*/Maşină n-are nici una
\hspace{1cm} cars/car not-he-has no one
\end{enumerate}

Additional evidence for the assumption that plural Num is the unmarked value for count nouns may be found in morphology: under the hypothesis that plural morphology on the noun head and the singular indefinite article are realizations of the Num head (see above), the plural Num is morphologically lighter, as expected for an unmarked value.

3.3. Adjective inflection mismatch

As illustrated in (6)) (repeated here), in German adjectives in the topic don’t show the “weak” inflection used when following a determiner, but the “strong” inflection used in the absence of a determiner bearing a strong ending:

\begin{enumerate}
\item a. Er hat keine polnischen Gänse  
\hspace{1cm} German (Fanselow 1988)
\item b. Polnische Gänse hat er keine  
\hspace{1cm} Polish-e geese has he no
\end{enumerate}

We explain this phenomenon using the same concept of morphological agreement by which we have described number mismatch: the establishment of the strong vs. weak form of the adjective is done after movement and is limited to the phase domain - exactly the portion of the structure which, under phase theoretical assumptions, is sent to the PF interface immediately after the valuation of all of its unvalued assumptions.

3.4. Remark on “PF-agreement”

We conclude that if the movement analysis is correct, then number mismatch and adjective inflection mismatch show that in some cases DP-internal agreement does not reach the moved part of the DP. We explained this fact by assuming that (sometimes) DP-internal agreement operates only on phase domains, by the time the phase is completed and the domain becomes opaque, so that the moved phrase, which by this time already occupies the edge of the phase or has moved even further, should not be affected.

Agreement realized only on phase domains, by the time the phase is completed and the phase domain becomes opaque, is likely to be a part of the PF derivation (on post-syntactic agreement, \textsuperscript{6} In Romanian a restricted class of verbs allow bare singular objects: verbs related somehow to possession, such as have, get, buy, bring, search, wear. The bare singular is interpreted as a “semantic incorporated nominal” (see Farkas - de Swart 2003).
see also Bobaljik 2005\(^7\)). This implies a more syntactic PF, which is in accordance with the idea of a cyclic Spell-Out. Operations at this level are characterized by the following properties:
- they read the output of the syntactic derivation
- they do not feed further syntactic derivation
- they have no interpretive effects
- they are local, limited to a Spell-out domain - in our case, the phase domain of a DP.

For the technical implementation of this kind of agreement, we have at our disposal the tools developed in the Distributed Morphology framework, such as the concept of *dissociated morphemes* (morphemes inserted *post-syntactically* in certain syntactic configurations, see Embick 1997).

If this kind of agreement should be kept as a part of (narrow) syntax, we could adopt Chomsky’s (2005) suggestion that various syntactic processes may take place in parallel inside a phase. In this case, valuation of uninterpretable features on Ns and As could take place inside the phase domain, triggered by the phase head D. Independently, the edge feature of this head could trigger raising of the NumP to SpecDP, thus escaping the phase-domain-agreement.

### 3.5. Lack of *de* insertion with complex numerals in Romanian

We are left with the problem raised by the contrast in (7):

<table>
<thead>
<tr>
<th></th>
<th>(7)</th>
<th>Romanian</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Au venit douăzeci de copii</td>
<td>have come twenty of children</td>
</tr>
<tr>
<td>b.</td>
<td>(*De) copii au venit douăzeci (*de)</td>
<td>(of) children have come twenty (of)</td>
</tr>
<tr>
<td>c.</td>
<td><em>Au venit douăzeci</em> copii</td>
<td>have come twenty children</td>
</tr>
</tbody>
</table>

The important fact to be noticed here is that *de* is not a true preposition, but a spell-out of some functional head in the nominal projection. As shown in (37), *de* is inserted only after (complex) cardinals which bear their own ϕ-features, which makes it look like a K head, spelling out a genitive case assigned to the noun by the numeral. However, (38) shows that the numeral is not a nominal head in this structure: if a determiner appears above the cardinal, it agrees with the lexical head and not with the cardinal.

<table>
<thead>
<tr>
<th></th>
<th>(37)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>două zeci / două sute de copii</td>
<td>two.fem tens / two.fem hundred of children (masc)</td>
</tr>
<tr>
<td>b.</td>
<td>doi / cinci / unsprezece (*de) copii</td>
<td>two.masc / five / eleven (one-sprezece) children</td>
</tr>
<tr>
<td>(38)</td>
<td>cei /<em>cele două zeci</em> de copii</td>
<td>the.masc.pl./the.fem.pl. two.fem tens of children (masc)</td>
</tr>
</tbody>
</table>

---

\(^7\) Bobaljik 2005 proposes a different type of argument for post-syntactic agreement: “if agreement is dependent on the outcome of a post-syntactic operation, then agreement must also be post-syntactic.” He applies this argument to verbal agreement, showing that it is depending on m-case, not on grammatical function; he relies on the idea that m-case itself is post-syntactically established.
So we infer that the cardinal occupies the Spec of a functional projection below D, let’s say Q₈; some cardinals still have an N feature, so they absorb Case, and trigger de insertion in order to case-license the NumP. de could be assumed to sit in Q₈.

In the topic construction, the NumP complement of this Q₀ moves. As a prepositional-like element, Q₀ is phonetically cliticized onto its complement. It follows that when the complement is phonetically null, Q₀, lacking a host, will also be null.

Notice that the split construction is also possible in DPs of the form Quantitative noun - de - Lexical Noun, where agreement on the verb shows that the lexical head of the construction is the second noun, so that the quantitative noun occupies a Spec position in the functional domain of the second noun:

(39)  
   a. O mulțime de studenți au / *a căzut la examen
       a fool of students have/has failed the exam  
   b. Studenți au căzut la examen o mulțime
       students have failed the exam a fool

With group nouns, where the second noun is not the lexical head of the DP, the split is impossible:

(40)  
   a. Un grup de fete a( ?u) cerut o reexaminare
       a group of girls has/have a re-examination  
   b. * Fete a(u) cerut o reexaminare un grup
       a group of girls has/have a re-examination

This contrast confirms the idea that the de which is dropped in the split-construction is the spell-out of a functional head in the extended nominal projection. (As shown in the example under note 3, pg. 10, the same phenomenon is found with mass nouns introduced by a quantity noun).

We may find other instances of deletion of a dummy functional head before a null element, including cases when the null element is the result of movement: in Welsh (see Rouveret 2006), in the construction copula + progressive particle + VP, the progressive particle disappears when the VP is fronted:

(41)  
   a. Yr ydw i’n teimlo’n chwyrn
       Prt am I Prog feel Pred agitated “I am feeling agitated”  
   b. Teimlo’n chwyrn yr ydw i
       feel Pred agitated Prt am I

References

---

8 I assume that cardinals generally occupy Spec positions because they may be phrasal:

(i)  
   a. între trei şi cinci ore
       between three and five ours  
   b. mai puţin de şapte metri
       more little of five meters
Frey, W. (2005) “Pragmatic properties of certain German and English left peripheral constructions”, Linguistics 43/1, 89-129