Subnominal empty categories as subordinate topics

PETRA SLEEMAN

0. Introduction

It is a well-known fact that in languages such as French there is a difference between subjects and objects w.r.t. noun ellipsis.[1] Whereas the noun can be left out from an indefinite noun phrase in subject position in (1), this is not possible if it is in object position, as in (2):

(1)   Trois parlent l'italien.
          'Three speak Italian.'

(2)   * J'ai vu trois.
          'I have seen three.'

In French, the noun can only be left out in indefinite objects if an overt quantitative pronoun is used:

(3)   J'en ai vu trois.
          'I have seen three of them.'

Interestingly, not only ellipted nouns as in (1)-(2) present a subject-object asymmetry, but the quantitative pronoun *en* as well, but this time the asymmetry is the reverse of the one in (1)-(2): whereas the quantitative pronoun *en* can be combined with an indefinite object (3), this is not possible with a subject, even not with the subject of an unaccusative verb:

(4)   * Trois en ont été lus par Paul.
          'Three have been read by Paul.'

In this respect the quantitative pronoun *en* differs from the genitive pronoun *en*, which does not present the same asymmetry:

(5)   Paul en a écrit la préface.
          'Paul has written the preface (of it).'
(6)   La préface en a été écrite par Paul.
          'The preface (of) it has been written by Paul.'

In the generative literature several syntactic accounts of the subject-object asymmetries in (1)-(4) or the absence of an asymmetry in (5)-(6) have been proposed, e.g. Ruwet (1972), Belletti & Rizzi (1981), Pollock (1986), Rizzi (1990), Cardinaletti & Giusti (1991). I will only discuss a more recent proposal, the one in Pollock (1998). I will show that there are data that are problematic for this syntactic approach. Although I assume that in ungrammatical

---

sentences with *en* the ungrammaticality is in some cases the result of the movement of *en*, for other cases I will explore whether a pragmatic/semantic account can be given for *en* being used or not, more specifically if this can follow from the discourse status of the nounless DP as a topic or a focus (also called a 'comment').

The paper is organized as follows. In 1, Pollock's account of noun ellipsis is discussed. In 2, I propose a pragmatic/semantic account of the noun ellipsis asymmetry. In 3, I also propose a pragmatic/semantic explanation for the quantitative pronoun *en* asymmetry, although I show that the use of *en*, which is a clitic, may sometimes also be blocked for syntactic reasons. In 4, the quantitative pronoun is compared to the genitive pronoun *en*. Finally, in 5, I summarize the results.

1. A syntactic account of noun ellipsis

According to Pollock (1998) there is a relation between the ungrammaticality of (7) and that of (8):

(7) * J'ai déjà expédié deux *pro* à Anne.
        'I have already sent two to Anne.'
(8) * les livres dont j'ai déjà expédié deux *pro* à Anne ...
        'the books of which I have already sent two to Anne ...'

In Pollock's analysis, the lexico-semantic features of *pro* are licensed ("recovered" locally from) those of *dont*. *Dont* and *deux pro* form a CP, inside a DP. *Deux pro* moves to Spec,DP:

(9) $_{DP}[[$deux *pro*$]_i$ F° $_{CP}[$dont $t_i]]$

Dont has to move as a head to C^o in the initial field of the relative clause. However in (8) the movement to C^o is blocked by the intervening V and I. Since *dont* is not a clitic it cannot move to C^o by adjoining to V and I. In (10), however, *dont* is extracted from the subject-DP. Since there are no intervening heads in this case, *dont* can move to C^o. The result is grammatical:

(10)   les livres $_{C^o}[$dont$_i]$ trois t_i *pro* viennent d'être vendus ...
         'the books of which three have just been sold ...'

Pollock assumes that if *dont* is the complement of a lexical noun instead of the complement of an empty head, it is phrasal instead of a head. In that case there are no blocking effects if *dont* is extracted from a DP in complement position. That is why (11) is grammatical:

(11)   le livre $_{CP}[$dont$_i$ $_{C^o}[_{IP}[$j'ai lu t_i le premier chapitre]]]$ ...
         'the book of which I have read the first chapter'

For (7) Pollock proposes that there is a null topic operator instead of *dont*, which, just like *dont*, is a head in French, and which also licenses *pro*. Since head movement of the empty operator is also blocked by the intervening V and I, (7) is ungrammatical. If the empty operator is extracted from subject position, there are no blocking heads, and the empty

operator can move to C°:

(12) $_{C°}[Op_i]$ t_i Trois *pro* viennent d'être vendus.
         'Three have just been sold'

Quantitative *en* also licenses *pro*. It is a head and attaches to a verbal host. There are no blocking heads in (13):

(13)   J'en$_i$ ai lu t_i trois *pro*.
         'I have read three (of them).'

Quantitative *en* can only license *pro* if it minimally c-commands *pro* at spell-out. This explains the contrast between (14) and (15). In (15), the lexical noun does not have to be licensed by *en*:

(14)   * [t_i Trois *pro*] en$_i$ seront publiés demain.
         'Three (of them) will be published tomorrow.'
(15)   [t_i Le premier chapitre] en$_i$ sera publié demain.
         'The first chapter of it will be published tomorrow.'

   Although for this restrictive set of data Pollock's analysis seems to be on the right track, it runs into problems if we add other data, e.g. the following ones:

(16)   ses livres dont$_i$ j'ai lu t_i le premier *pro*
         'his books of which I have read the first'
(17) Op_i J'ai lu t_i le premier *pro*.
         'I have read the first.'
(18)   * J'en$_i$ ai lu t_i le premier *pro*, de livre.
         'I have read the first of them (= the first book).'
(19)   J'ai lu $_{DP}[[$deux *pro*$]_i$ F° $_{CP}[$de ses livres $t_i]]$
         'I have read two of his books.'

Although *dont* and the null topic operator license *pro*, (16)-(17) should be ungrammatical, because, in Pollock's analysis, if they license *pro*, they are heads. A phrasal analysis such as Pollock proposes for *dont* with a lexical noun in (11), is not possible for *dont* in combination with *pro*, because Pollock assumes that phrasal *dont* is featureless and would therefore fail to license *pro*. The same reasoning would apply to the empty operator. However, if *dont* in (16) and the null topic operator in (17) are heads, it is not clear why these sentences are grammatical. Just as in (7)-(8), V and I block movement to C°. As for (18), its ungrammaticality cannot be explained. Just as in (13), *pro* is licensed by *en* and movement of *en* is not blocked. In (19), *de ses livres* does not minimally c-command *pro* at spell-out, so that *pro* is not licensed.
   Because of these problems, I will propose another explanation for the subject-object asymmetries in French. What I would like to propose is that there is not a purely syntactic reason for the subject-object asymmetries, but that there is rather a pragmatic/semantic reason: the nounless DP has to be semantically related to a DP in the linguistic context, which

is only possible if the relation is feasible from a pragmatic point of view.

2. A pragmatic/semantic account of the noun ellipsis asymmetry

If the noun is left out of a DP, a relation has to be established with the linguistic context in order to provide the nounless gap with an interpretation. I propose that this semantic relation can only be established if the empty noun can qualify as a topic, although part of the DP can qualify as a focus. I illustrate this by means of Erteschik-Shir's (1997, 1999) theory of focus structure.

Erteschik-Shir (1997, 1999) tries to account for relations between sentences by adding the level of Focus Structure to the grammar, a level that mediates between syntax and semantics and that feeds PF because it provides the explicit phonetic spell-out including intonation. Focus structure (f-structure) is an annotated structural description in which Topic and Focus constituents are marked.

Erteschik-Shir's f-structure theoretical approach to interpretation is dynamic in that it assumes a theory of discourse which defines the state of the common ground both before and after the utterance of the sentence. The common ground consists of a set of file cards with indexed headings which represent existing discourse referents. The common ground propositions form entries on these file cards. The cards on top of the file are licensed as potential topics of an utterance.

Several f-structure rules apply to f-structures (structural descriptions annotated for Topic and Focus), among which the following:
- TOPIC instructs the hearer to locate on the top of his file an existing card.
- FOCUS instructs the hearer to either open a new card and put it on top of his file (in the case of an indefinite) or locate an existing card and put it on the top of the file (in the case of a definite).

When a hearer hears the following sentence, in which *I* is licensed as the topic, he selects the corresponding card from the top of the file and evaluates the rest of the sentence w.r.t. the topic:

(20)   I   have a dog.
        TOP   FOC

He also opens a new card, which he labels *dog* and which he puts on top of the file. Both *I* and *dog* are now on top of the file and available as future topics. When the speaker adds the following sentence, the hearer selects *dog* as the new topic, because of the form of the pronoun:

(21)   It   is brown.
        TOP FOC

Besides individual topics such as *I* or *dog*, there are also stage topics. Stage topics may be overt (*this afternoon*, *on Park Avenue*) or discoursally implied (the here-and-now). In (22), the whole sentence is taken as a focused event predicated of a covert stage topic (the here-and-now):

4

(22)   It is raining.
            FOC

F-structures can also be assigned to subconstituents of the utterance. They are called subordinate f-structures. An example of a construction involving a subordinate f-structure is the partitive construction. The subordinate f-structure is formed around a card on top of the file which represents a discoursally available set. A constituent which defines a subset of this topic set is focused triggering the partitioning of this set:

(23)   I [met [two]$_{\text{FOC-sub}}$ [of the students]$_{\text{TOP-sub}}$]$_{\text{FOC}}$

The new subset card is now available as the main topic:

(24)   [[Two]$_{\text{FOC-sub}}$ [of the students]$_{\text{TOP-sub}}$]$_{\text{TOP}}$ [are intelligent]$_{\text{FOC}}$

The assignment of subordinate f-structure also explains the following contrast:

(25) * [A friend]$_{\text{TOP}}$ is intelligent.
(26)   [[A friend]$_{\text{FOC-sub}}$ [of mine]$_{\text{TOP-sub}}$]$_{\text{TOP}}$ is intelligent.

In (25) the indefinite DP has neither a generic nor a contrastive reading and cannot qualify as the topic. The indefinite DP in (26), however, has a specific interpretation and therefore can function as a topic. The utterance (26) instructs the hearer to open a new card for *a friend*, the subordinate focus, whereas the subordinate topic *I* already exists as a card. Once this card is opened, the indefinite is (speaker) referential and qualifies for topic status. This explains the grammaticality of (26).

F-structures are constrained by the Topic Constraint, which says that the unmarked f-structure is one in which either the subject or a stage is the topic. This explains the ungrammaticality of (25), in which the indefinite cannot function as a topic with the individual level predicate.

The f-structure in the following sentence takes the subject pronoun as the main topic and the object pronoun as the subordinate topic:

(27)   [He]$_{\text{TOP}}$ [criticized [her]$_{\text{TOP-sub}}$]$_{\text{FOC}}$.

How can a discourse theory such as Erteschik-Shir's Focus Theory account for the noun ellipsis data? Since ellipted nouns have to be known from the context to receive an interpretation, their referent must be available as a card on top of the file, i.e. must have been introduced in the domain of discourse. Ellipted nouns are thus topics. They cannot be newly introduced individuals. We have seen, however, that there are main topics and subordinate topics. In the partitive construction (28), *de mes collègues* is a subordinate topic whereas *deux* is the subordinate focus:

(28)   [Il]$_{\text{TOP}}$ [a insulté [deux]$_{\text{FOC-sub}}$ [de mes collègues]$_{\text{TOP-sub}}$]$_{\text{FOC}}$.
            'He has insulted two of my colleagues.'

Now, I claim that nouns can only be ellipted either if the whole DP is a topic, as in (29):

(29)   Il a attrapé trois lions. [Il]$_{TOP}$ a tué [les trois pro]$_{TOP\text{-}sub}$.
        'He has caught three lions. He has killed all three.'

or if the ellipted noun is a subordinate topic, with the quantifier functioning as the subordinate focus:

(30)   [[Deux]$_{FOC\text{-}sub}$ [pro]$_{TOP\text{-}sub}$]$_{TOP}$ sont intelligents.
        'Two are intelligent.'

Since according to the Topic Constraint, subjects tend to be topics, there will be a subordinate f-structure with a focalized quantifier and the noun as a subordinate topic in indefinite subject DPs. Since the noun is a topic, it must be discoursally available. The noun can be left out in (30), because the subordinate f-structure implies that it is a topic so that there must be an available card on top of the file.

In the analysis in which partitives involve an empty noun, the empty noun is a subordinate topic, because it is coreferentiel with the noun within the partitive PP, which is a topic:

(31)   [Il]$_{TOP}$ [a insulté [deux]$_{FOC\text{-}sub}$ [[pro] $_{TOP\text{-}sub}$ de mes collègues]$_{TOP\text{-}sub}$]$_{FOC}$.
        'He has insulted two of my colleagues.'

The use of a definite determiner before the ordinal in (32) also implies the definition of a restrictive set. What is asserted is that from a discoursally available set I have bought the one that is third.

(32)   J'[ai acheté [le troisième]$_{FOC\text{-}sub}$ [pro]$_{TOP\text{-}sub}$]]$_{FOC}$
        'I have bought the third.'

That ellipted nouns are always topics is confirmed by the fact that a *de* N-dislocation is possible in combination with an empty noun. Dislocated constituents are always topics:

(33)   Je préfère la bleue *pro*, de robe.
        'I prefer the blue dress.'

Indefinite noun phrases in object position are generally interpreted as foci without a subordinate f-structure:

(34)   J'[ai lu trois livres]$_{FOC}$.
        'I have read three books.'

Sentence (35) is like (34) in that no subordinate f-structure is formed. Since the noun does not qualify as a subordinate topic, which must be discoursally available, it cannot be left out:

(35) * J'[ai lu trois *pro*]$_{FOC}$.
        'I have read three.'

This also holds for the associate in an impersonal construction:

(36)   Il est arrivé trois linguistes.
        'There arrived three linguists.'
(37) * Il est arrivé trois *pro*.
        'There arrived three.'

Since there is no trigger for the formation of subordinate f-structure in (37), the empty noun does not qualify as a subordinate topic.

It has been observed by Milner (1978) that with some weak quantifiers noun ellipsis in object position is more acceptable than with others. Milner gives the following examples with *dont*, which I analyze as a base-generated relative pronoun, following Godard (1988):

(38)   des livres dont j'ai lu plusieurs / certains / quelques-uns / un grand nombre / des
        tonnes / des centaines / *beaucoup / *trois
        'books of which I read several / certain ones / some / a lot / tons / hundreds / many /
         three'

The reason for the increased acceptability of noun ellipsis with *un grand nombre*, *des tonnes* and *des centaines* might be that the quantifier is a nominal expression, which can function itself as the nominal complement of the verb and is not necessarily the specifier of a noun. The reason for the increased acceptability of noun ellipsis with *plusieurs*, *certains* and *quelques-uns* might be that these quantifiers indicate a quantity that is evaluated w.r.t. a set, which may be undefined. C*ertains* and *quelques-uns* indicate a quantity that is negatively evaluated by the speaker, since it is small subset w.r.t. the, possibly undefined, whole set, *plusieurs* indicates a quantity that is positively evaluated, since it indicates more than the subset of one (see Gondret 1976). This implied partitive interpretation might lead to topicalization of the head noun so that it can be omitted even in object position.

My analysis also predicts that noun ellipsis is not possible in postverbal subjects that present hearer-new information, i.e. that are foci. One such case is the indefinite subject in the stylistic inversion sentence in (39):

(39) * le jour où sont arrivés trois.
        'the day that three arrived'

Kayne's (1986) account of the ungrammaticality of (39) is based on Belletti & Rizzi's (1981) theory of noun ellipsis. The postverbal subject containing PRO is governed by V, which is not allowed. In my analysis, the ungrammaticality of (39) is the result of the whole postverbal constituent being focalized. If subordinate-f structure is formed, i.e. if only part of the postverbal subject is focalized and the empty noun is a topic, the result is grammatical:

(40)   le jour où est arrivé [[le troisième]FOC-sub [*pro*]TOP-sub]FOC
        'the day that the third one arrived'

In this section I have argued that subordinate f-structure, which makes a semantic relation

between the nounless DP and a DP in the context possible, is in principle not possible in an indefinite object introduced by a weak quantifier. This restriction on Focus Structure follows from the Topic Constraint, which says that the unmarked f-structure is one in which either the subject or a stage is the topic. Objects tend to be foci, unless a special device is used, such as the pronoun *her* in (27), in which case the object can be a subordinate topic. As a consequence of the Topic Constraint, subordinate topics with a focalized quantifier are natural in subject position but not in object position, unless a special device is used such as a partitive PP or a definite determiner, which mark the (empty) noun within the DP as a (subordinate) topic. In the next section I will argue that the quantitative pronoun *en* is another device to create a subordinate topic. I will show, however, that the use of *en* is restricted because of syntactic considerations.

3. Quantitative *en*

The ungrammatical sentences (35) and (37) become grammatical if the quantitative clitic *en* is added:

(41)   J'en ai lu trois.
        'I have read three of them.'
(42)   Il en est arrivé trois.
        'There arrived three of them.'

However, the use of *en* is not possible in combination with a definite DP in object position, which is problem for a syntactic analysis such as Pollock's (1998), see section 1:

(43) * J'en ai lu le troisième.
        'I have read the third one of them.'

The ungrammaticality of (43) must be due to the presence of the definite determiner, because (44), with an indefinite determiner, is grammatical:

(44)   J'en ai lu un troisième.
        'I have read a third one.'

The contrast between (43)-(44) has generally been accounted for in syntactic terms in the generative literature. Sleeman (1992) assumes that the definite determiner in (43) is a head, whereas the indefinite determiner in (44) is a specifier. The contrast in grammaticality is attributed to Relativized Minimality (Rizzi 1990). The trace of *en*, which is assumed to move as a head, cannot be antecedent-governed by *en* in (43) because the filled D° functions as a potential antecedent governor. The indefinite determiner in (44) is not a head, but a specifier, and, as a result, does not block antecedent-government of the trace by *en*.

In Vermandere's (2002) analysis, *en* is doubled by *pro*, which has nominal features. Movement of *pro* to the clitic position in (43) is blocked because the filled D° intervenes between the *en*-clitic and *pro*. The definite determiner spells out a nominal feature and (default or marked) φ-features. Since this φ-feature composition is identical to the one *pro* is

endowed with, movement of *pro* is barred: the determiner itself is closer to the *en*-clitic. The derivation eventually crashes with a 'frozen' *pro*.

    A problem for both analyses of the ungrammatical (43) is the contrast between (45)-(46):

(45) * Elle en salua chacun.
         'She greeted every one of them.'
(46)   Elle en salua un seulement.
         'She only greeted one of them.'

To explain the ungrammaticality of (45), it would have to be assumed that *chacun* is a head, just like the definite article, wheras *un* is a specifier or that *chacun*, but not *un*, has the same φ-features as the definite article. Both solutions seem rather ad hoc. Another possibility would be that *un* is moved from an N position to adjoin to the specifier *chaque* 'each' (for such an analysis see Junker & Vinet 1993). If there is already an N, *un*, there is no room for quantitative *en*, which is also generally assumed to move from the N or NP position. However, such an analysis could then also be adopted for *quelques-uns* 'some', which is the pronominalized form of *quelques*. With *quelques-uns*, however, *en* can be used:

(47)   J'en ai lu quelques-uns.
         'I have read some (of them).'

I therefore assume that *chacun* and *quelques-uns* are taken as morphological units from the lexicon and that they modify a base-generated empty noun or the trace of *en*, in the case of *quelques-uns*. This analysis is supported by the fact that until the 16th century *chacun* could be used with an overt noun (Junker & Vinet 1993):

(48)   Entre chascune tour estoit espace de troys cens douze pas. (Rabelais)
         'Between every tower there was a space of 312 steps.'

    In Sleeman (1996) I gave a semantic account of the ungrammaticality of (43) and (45). Since the definite determiner and *chacun* are "strong" in the sense of Milsark (1974), the nounless DP they are part of can be related to a DP in the discourse without the use of *en* being necessary. A derivation without *en* is more economical and therefore the preferred one:

(49)   J'ai lu le troisième.
         'I have read the third one.'
(50)   Elle salua chacun.
         'She greeted each one.'

    In this paper I will pursue this idea. In the previous section, I claimed that a semantic relation between a nounless DP and a DP in the context is only possible if the nounless DP is a topic or if it has subordinate f-structure, i.e. if the empty noun is a topic whereas the quantifier is focalized. It follows from the Topic Constraint that the unmarked f-structure is one in which either the subject or a stage is the topic. Objects tend to be foci. A sentence such as (51), in which the whole object DP is a focus, is ungrammatical, because no relation with a DP in the context can be established to provide the empty noun with an interpretation:

(51) * J'ai lu trois.
      'I have read three.'

One of the devices to create subordinate f-structure is to add a partitive PP, see (52) repeated from (31):

(52)   [Il]$_{TOP}$ [a insulté [deux]$_{FOC\text{-}sub}$ [[*pro*] $_{TOP\text{-}sub}$ de mes collègues]$_{TOP\text{-}sub}$]$_{FOC}$.
          'He has insulted two of my colleagues.'

   The use of a definite determiner before the ordinal in (53), which is repeated from (32), also implies the definition of a restrictive set:

(53)   J'[ai acheté [le troisième]$_{FOC\text{-}sub}$ [*pro*]$_{TOP\text{-}sub}$]]$_{FOC}$
          'I have bought the third.'

   Another device to create subordinate f-structure is the use of the clitic *en*. If the quantitative pronoun *en* is used, which refers to a card on top of the file, partitioning of this topic set is possible by means of focalization of the quantifier so that subordinate f-structure is formed:

(54)   J'en$_i$ ai lu [[trois]$_{FOC\text{-}sub}$ [t_i]$_{TOP\text{-}sub}$]]$_{FOC}$.
          'I have read three of them.'
(55)   Il en est arrivé [[trois]$_{FOC\text{-}sub}$ [t_i]$_{TOP\text{-}sub}$]$_{FOC}$.
          'There arrived three of them.'

   It has to be noted that a topic set can be a restrictive set or a non-restrictive set itself. In (56) the topic set is a restrictive set consisting of four books, which is partitioned by means of the focalized quantifier to form a new set of three books, a subset of the first:

(56)   J'ai acheté quatre livres l'après-midi. J'en ai lu trois le soir même.
          'I bought four books in the afternoon. I read three of them the same evening.'

In (57) just the head on the topic card serves to form a new set:

(57)   Hier j'ai lu deux livres. Aujourd'hui j'en ai lu trois.
          'Yesterday I have read two books. Today I have read three.'

In both readings (51) is ungrammatical. I claim that this is so because in both cases subordinate f-structure, and hence a relation with a DP in the context, is not possible in the indefinite object introduced by a weak quantifier, unless the clitic *en* or a partitive PP is added to trigger subordinate f-structure.
   I claim that *en* is only used to create subordinate structure. If f-structure is created in another way, *en* is not used. This explains the ungrammaticality of *en* with a partitive PP (58) and with a definite DP or a DP introduced by a strong quantifier, cf. (43) and (45), repeated here as (59)-(60):

(58) * J'en ai lu trois de ses livres.
        'I have read three (of them) of his books.'
(59) * J'en ai lu le troisième.
        'I have read the third one of them.'
(60) * Elle en salua chacun.
        'She greeted each one of them.'

Dont cannot be combined with *en* either. Since *dont* cannot be combined with an indefinite object (but see (38) for some exceptions), which is a focus, we would expect *en* to be possible. For some speakers it is (see Hanse 1987). For those who do not accept it, the reason might be that a genitive phrase in a sentence initial position cannot be doubled by *en*:

(61) ?? Ces livres, dont j'en ai lu deux.
        'These books, of which I have read two (of them).
(62) ?? De ces livres, j'en ai lu deux.
        'Of these books, I have read two (of them).'
(63)   Ces livres, j'en ai lu deux.
        'These books, I have read two of them.'

Corblin (1995:125) observes that *en* can optionally be used in combination with a partitive PP if this PP contains itself an empty noun:

(64)   J'(en) ai pris dix des bleues *pro*.
        'I have taken ten of the blue ones.'

I suggested that in partitive constructions the empty noun is a subordinate topic, because it is coreferentiel with the noun within the partitive PP, which is a topic. Now, if the partitive PP contains itself a empty noun, the first empty noun can only indirectly get its reference. Therefore, its subordinate topic status is not so clear and *en* can be used to mark it as a topic:

(65)   [J']$_{TOP}$ [[en]$_{i\ TOP\text{-}sub}$ ai pris [dix]$_{FOC\text{-}sub}$ [[t]$_{i\ TOP\text{-}sub}$ des bleues *pro*]$_{TOP\text{-}sub}$]$_{FOC}$.

Since according to the Topic Constraint, preverbal subjects tend to be topics, subordinate f-structure is naturally formed within a nounless indefinite subject. Therefore the use of *en* is also superfluous in this case, which explains, in my view, the ungrammaticality of (4), repeated here as (66):

(66) * Trois t_i en$_i$ ont été lus par Paul.
        'Three of them have been read by Paul.'

In my view, *en* is the lexical variant of the empty pronominal *pro*. The function of *en* is to create subordinate f-structure. In this respect *en* differs from the partitive PP, which can also be used as an explicit topic marker even if the DP already has acquired subordinate f-structure by means of another device:

(67)   J'ai acheté le troisième de ses livres.
         'I have read the third one of his books.'
(68)   Trois de ses livres sont très intéressants.
         'Three of his books are very interesting.'

Since, in my analysis, *en* is the lexical variant of the empty pronominal *pro*, it always originates as the semantic head of the DP. In my view there is no distinction between quantitative and partitive *en* (see e.g. Milner 1982, who argues that the first one is a head, whereas the second one is the complement of an empty head). Pollock (1998) always analyzes *en* as partitive *en*. However, if *en* has the same function as a partitive PP, the combination with a nounless DP that already has subordinate f-structure should be possible (cf. (67)). However, it is not, which suggests that *en* is always quantitative:

(69) * J'en ai acheté le troisième, de ses livres.
         'I have bought the third one, of his books.'

Haïk (1982) observes that there are speakers for whom *en* is possible in (70) but not in (71):

(70)   Beaucoup en sont gâtées, de ces pommes.
         'Many of these apples are rotten.'
(71) * Beaucoup en sont gâtées, de pommes.
         'Many apples are rotten.'

These data suggest that for those speakers who accept (70), un uneconomical use of quantitative *en* is allowed in combination with a subject. I suggest that if (71) is felt to be worse, this might be so because the Topic Constraint makes a partitive interpretation in subject position more natural than a quantitative interpretation:

(72) ?? Beaucoup sont gâtées, de pommes.
(73)   Beaucoup sont gâtées, de ces pommes.

Although the reason for the use of quantitative *en* is thus a semantic one, in my analysis, all ungrammaticalities cannot be attributed to its uneconomical use. Since *en* is a clitic, it has to be moved out of the DP to a verbal host. This is why *en* cannot be used with a postverbal subject that presents hearer-new information.

(74) * le jour où en ont téléphoné trois.
         'the day that three (of them) telephoned.'

In Kayne & Pollock's (2001) view, such data suggest that the postverbal subject is originally in a high, subject-like, position. The clitic must be extracted to a position c-commanding its original position, which is possible if it is extracted from an object but not if it is extracted from a (high) subject position.

The combination of *en* with an indefinite PP is also grammatical from a semantic point of view, but is ruled out by the illegal movement of *en* out of the PP, see Kayne (1975, 1981):

(75) * Je m'en$_i$ suis pourvu $_{PP}$[de trois t_i].
        'I have provided myself with three (of them).'

In this respect quantitative *en* resembles genitive *en*:

(76) * J'en$_i$ ai parlé du premier chapitre t_i.
        'I have spoken about the first chapter of it.'

In the next section I will show, however, that in several other respects quantitative *en* differs from genitive *en* and I will argue that the proposed analysis of quantitative *en* can account for the differences.

4. Genitive *en*

Genitive *en* does not present a subject/object asymmetry as does quantitative *en*. It can be extracted from an object, just like quantitative *en*, see (77), but although it normally cannot be extracted from a subject (Ruwet 1972), see (78), it contrasts with quantitative *en* in that it can be combined with the subject of an unaccusative verb, cf. (79) with (66), repeated here as (80):

(77)   Il en$_i$ a lu le premier chapitre t_i.
          'He has read the first chapter of it.'
(78) * Le premier chapitre en$_i$ t_i traite de sa vie.
          'The first chapter (of it) deals with his life.'
(79)   La préface t_i en$_i$ a été écrite par Paul.
          'The preface (of it) has been written by Paul.'
(80) * Trois t_i en$_i$ ont été lus par Paul.
          'Three of them have been read by Paul.'

The explanation that is generally given for the grammaticality of (77) as opposed to the ungrammaticality of (78), is that in (78), *en* is moved to a non c-commanding position. In (79) the complement of the unaccusative verb moves to subject position after the extraction of *en* (Ruwet 1972):

(81) $_{DP}$[La préface t_i]$_j$ en$_i$ a été écrite t_j par Paul.

Since a derivation as in (81) is also possible for (80), I have attributed the ungrammaticality of this sentence to an economy violation. Since subordinate f-structure is created in subject position, *en* is not needed.
    It can be concluded from the contrast between (79) and (80) that the use of genitive *en* does not lead to an economy violation and therefore resembles the partitive PP (cf. (67)-(68)). I propose that this is so because the function of genitive *en* is not to create subordinate f-structure, but only to explicitly mark the relation with a DP in the discourse. Therefore it can also be used in combination with a definite DP. The use of genitive *en* is not necessary. Both

with a subject and with an object it can be omitted, although Azoulay (1979) observes that with an object it is more required than with a subject. This must be due to the Topic Constraint. Subjects tend to be topics and therefore tend to be more easily in relation with the discourse, whereas objects tend to be foci, i.e. new information:

(82)   La préface a été écrite par Jean.
         'The preface has been written by Paul.'
(83) ? Paul a écrit la préface.
         'Paul has written the preface.'

    That genitive *en* has another function than quantitative *en* is also suggested by the fact that genitive *en* is possible in (84), whereas quantitative *en* is not in (85), cf. (59):

(84)   J'en ai lu le troisième, de ce livre.
         'I have read the third (chapter), of this book.'
(85) * J'en ai lu le troisième, de chapitre.
         'I have read the third chapter.'

In syntactic analyses, it is difficult to account for this difference in grammaticality. In the previous section, I rejected some syntactic analyses of the ungrammaticality of (85) and proposed that it is due to an economy violation: the use of the quantitative pronoun *en* is superfluous because subordinate f-structure is formed otherwise. The use of genitive *en* in (84) does not lead to ungrammaticality, because its sole function is to mark the relation with the context.

5. Conclusion

In this paper I have proposed that subject-object asymmetries w.r.t. empty nouns receive an explanation in the pragmatic/semantic component of the grammar and not in the syntactic component. This was supported by empirical motivation: I showed that a syntactic analysis such as e.g. Pollock's (1998) can only account for a restricted set of data. I have rejected any syntactic analysis of the subject-object asymmetry and I have claimed that the asymmetries are the consequence of a constraint on the assignment of information structure to the output of the syntactic component. This constraint blocks the natural assignment of subordinate f-structure - with the noun as a subordinate topic and the quantifier as a subordinate focus - to DPs in object position, unless the empty noun can be licensed as a (subordinate) topic by lexical or syntactic topic markers. Only if the empty noun is a topic, it can receive its interpretation from a noun in the context.

References:

Azoulay-Vicente, Avigail (1979). "Article défini et relations anaphoriques en français." *Recherches Linguistiques de Vincennes* VII, 5-46.
Belletti, Adriana & Luigi Rizzi (1981). "The syntax of 'ne': some theoretical implications". *The Linguistic Review* 1, 117-54.
Cardinaletti, Anna & Giuliana Giusti (1991). "Partitive *ne* and the QP-hypothesis: A case study". *Working*

Papers in Linguistics, University of Venice.

Corblin, Francis (1995). *Les Formes de Reprise dans le Discours. Anaphores et Chaînes de Référence*. Presses Universitaires de Rennes.

Erteschik-Shir, Nomi (1997). *The Dynamics of Focus Structure*. Cambridge University Press, Cambridge.

Erteschik-Shir, Nomi (1999). "Focus Structure and scope." In G. Rebuschi & L. Tuller (eds.), *The Grammar of Focus*, Benjamins, Amsterdam & Philadelphia, 119-150.

Godard, Danièle (1988). *La Syntaxe des Relatives en Français*. CNRS, Paris.

Gondret, Pierre (1976). "*Quelques*, *plusieurs*, *certains*, *divers*: étude sémantique". *Le Français Moderne* 44, 143-152.

Haïk, Isabelle (1982). "On clitic *en* in French". *Journal of Linguistic Research* 2.1, 63-87.

Hanse, Joseph (1987). *Nouveau Dictionnaire des Difficultés du Français Moderne*. Duculot, Paris / Gembloux.

Junker, Marie-Odile & Marie-Thérèse Vinet (1993), "*Chaque*, quantifieur binominal: représentation d'un phénomène de variation interdialectale". *Travaux de Linguistique* 26, 5-20.

Kayne, Richard (1975). *Syntaxe du Français. Le Cycle Transformationnel*. Editions du Seuil, Paris.

Kayne, Richard (1981). "ECP extensions". *Linguistic Inquiry* 12, 93-133.

Kayne, Richard (1986). "Connexité et inversion du sujet". In M. Ronat & D. Couquaux (eds.), *La Grammaire Modulaire*, Editions de Minuit, Paris, 127-147.

Kayne, Richard & Jean-Yves Pollock (2001). "New thoughts on stylistic inversion". In A. Hulk & J.-Y. Pollock (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*, Oxford University Press, Oxford / New York, 107-162.

Milner, Jean-Claude (1978). *De la Syntaxe à l'Interprétation*. Seuil, Paris.

Milsark, Greg (1974). *Existential Sentences in English*. Dissertation, MIT.

Pollock, Jean-Yves (1986). "Sur la syntaxe de *en* et le paramètre du sujet nul". In M. Ronat & D. Couquaux (eds.), *La Grammaire Modulaire*, Editions de Minuit, Paris, 211-245.

Pollock, Jean-Yves (1998). "On the syntax of subnominal clitics: cliticization and ellipsis". *Syntax* 1:3, 300-330.

Rizzi, Luigi (1990). *Relativized Minimality*. MIT Press, Cambridge, Massachusetts.

Ruwet, Nicolas (1972). *Théorie Syntaxique et Syntaxe du Français*. Editions du Seuil, Paris.

Sleeman, Petra (1992). "Extraction out of DP in French and proper head government". In Reineke Bok-Bennema & Roeland van Hout (eds.), *Linguistics in the Netherlands 1992*, John Benjamins, Amsterdam / Philadelphia, 235-246.

Sleeman, Petra (1996). *Licensing Empty Nouns in French*. Holland Academic Graphics, The Hague.

Vermandere, Dieter (2002), "A unified analysis of French and Italian *en/ne*". In T. Satterfield, C. Tortora and D. Cresti (eds.), *Current Issues in Romance Languages: Selected Papers from the XXIXth Linguistic Symposium on Romance Languages (LSRL)*, Benjamins, Amsterdam / Philadelphia.