Abstract

In this paper, we present an analysis of noun phrases with elided nouns that dispenses with the positing of empty categories and preserves the NP structure assumed for NPs with overt nouns, modulo the absence of the head noun. On a par with traceless analyses of long distance dependencies, this is proposed as a further step towards a more lean theory of grammar, without phonetically null items.

1   Introduction

Elliptical NPs get structured around missing head nouns, as illustrated in the following examples from English (Lobeck (1995)):

(1) a. Although John's friends were late to the rally, [Mary's -] arrived on time.

   b. Because [her two -] were sick, Melissa didn't take the children to swimming lessons that week.

   c. We tasted many wines, and I thought that [some -] were extremely dry.

This is a widespread type of construction that occurs in many languages other than English, as exemplified below with data from German (Netter (1996)) and Spanish (Ticio (2005)):

(2) a. das rote Auto und [das blaue -]
      the red car   and   the blue
      the red car and the blue one

   b. la casa azul y   [la - verde]
      the house blue and   the   green
      the blue house and the green one

The following is a list of typical properties of NP ellipsis that have been reported in the literature.

As noun ellipsis is to be viewed as a phenomenon different from null arguments, at least one specifier, one complement or one modifier is present in the elliptical NP.

In some languages, like German, ellipsis cannot be NP initial (Netter (1996)):

(3) a. Alte Männer mit Hut haben [junge - mit Mütze] getroffen.
      old   men   with hat have   young with cap   met
      Old men in hats met young ones in caps.

---

b. * Alte Männer mit Hut haben [- mit Mütze] getroffen.

In some languages, some determiners, like the English definite articles, cannot alone form an NP (English example from Lobeck (1995)), while others are allowed to (1c):[1]

(4) * A single protester attended the rally because [the -] apparently felt it was important.

In languages like Portuguese and Spanish, with pre- and post-head adjectives, pre-head ones (which are intensional) cannot appear in this construction (Spanish example in (5b) from Ticio (2005)), although postnominal adjectives (intensional or not) can (Portuguese example in (5a)):

(5) a. a   terrorista real e   [a   - imaginada]
       the terrorist   real and   the imagined
       the real terrorist and the imagined one

    b. * Ayer   vi   a la verdadera terrorista y   a [la supuesta -].
         yesterday I saw the true   terrorist and   the alleged
         intended: *Yesterday I saw the real terrorist and the alleged one.*

In addition, the elliptical NP relies on an antecedent to be interpreted, from which it inherits gender as well as subcategorization and count/mass properties (Netter (1996) and Masullo (1999)),

(6) a. die starke Konzentration auf die Wirtschaft
       the strong concentration on the economy

       und [die weniger grosse - auf den Umweltschutz]
       and   the less   large   on the environment
       the strong concentration on the economy and the less large on the environment

    b. * Juan visitó a sus tíos   y   Pedro visitó a [la - suya].
         Juan visited his uncles/aunts.MASC and Pedro visited   the his.FEM
         intended: *Juan visited his aunt and uncle and Pedro visited his (aunt).*

but not necessarily number:

(7) Juan visitó a sus tíos   y   Pedro visitó a[l - suyo].
    Juan visited his uncles/aunts.MASC and Pedro visited   the - his.MASC.SG
    Juan visited his aunt and uncle and Pedro visited his (uncle).

---

[1]We are assuming, like much of the literature on noun ellipsis, that if an item can appear in an NP which is restrictively modified, it is not a pronoun but a determiner, since restrictive modifiers attach to $\overline{\text{N}}$ (see Section 4.4).

In English, an overt element (*one*) is used instead of a null noun in certain contexts (*one* anaphora).

Finally, NP ellipsis should not be confused with missing-N generics (e.g. *the desperately poor*),[2] which tend to be limited to descriptions of people and do not resort to an antecedent to be interpreted.

2   Previous Accounts

Many previous analyses of NP ellipsis, either in the HPSG framework (e.g. Netter (1996), Nerbonne and Mullen (2000)) or under other theoretical persuasions (e.g. Lobeck (1995), Ticio (2005)), assume an empty category approach where the missing noun is assumed to be an actual, though phonetically null, lexical item.

In line with a view of grammar free of reified empty categories, alternatives to this approach have been advanced as well. One of such alternatives was put forward in (Winhart, 1997) and consists in analyzing adjectives in elliptical NPs as the result of a nominalization lexical rule. A major problem for this account, pointed out in (Netter, 1996), is that it cannot derive an elliptical NP where the adjective has modifiers or specifiers of degree, as in (8).

(8) die ziemlich alten Männer und [die [besonders   jungen] -]
    the quite   old   men   and the particularly young
    the quite old men and the particularly young ones

A similar analysis, based on explaining away the data via some category change of the elements occurring in elliptical NPs, might be envisaged for determiners: when items from these categories appear in elliptical NPs, they could be taken as pronouns, either as a result of some lexical rule, or even as homonymous items included in the lexicon from the start. Such an approach has also found appropriate appreciation and criticisms in (Nerbonne and Mullen, 2000), the main argument against it being the possibility of restrictive modification (see Section 4.4).

Another line of research has been to propose the underspecification of adjectives and other NP elements so that they can function as nouns as well. A crucial problem here concerns how the semantics of the NP is composed given that determiners and nouns, for instance, make different contributions to its semantic content. This is the approach explored in (Beavers, 2003) for nouns and determiners. That work is limited in its range because it only covers elliptical NPs with a single determiner.

Another option to be explored for an analysis that does not resort to empty categories is to use a unary syntactic rule, which can operate in tandem with the usual specifier-head or adjunct-head schemata. This possibility is appreciated in (Netter, 1996), to be dismissed as being theoretically uninteresting. Taking into account,

---

[2]We will call missing-N generics to what is referred to in the literature as *people* deletion (Pullum (1975)) or null-N generics (Nerbonne and Mullen (2000)), because that expression is more neutral than the latter with respect to the status of transformations or empty categories.

however, how the use of unary schemata has been enhanced since then,[3] this is clearly an option worth considering, and it is the approach that will be explored in the next Sections.

Two computational HPSGs for German (Müller and Kasper (2000) and Müller (1996)) indeed use unary syntactic rules that apply to noun modifiers and produce a noun-headed projection.

The analysis proposed in the following Sections presents a unified treatment of noun adjuncts and determiners in noun ellipsis constructions, merging the latter accounts with the account of Beavers (2003).

3 Functors and NP structure

Before entering into the details of the proposed analysis for elliptical NPs, it is useful to briefly sketch the NP organization assumed by our analysis.

Our account of ellipsis builds upon the work of Allegranza (1998a), Allegranza (1998b), Van Eynde (2003a) and Van Eynde (2003b), who propose the simplification resulting from replacing the specific constructs used to handle specifiers and adjuncts by a more general one for functors.[4] Following this work, the specifier vs. adjuncts distinction becomes useless by letting all functors select their head via a single feature (its designation has not been uniform: here we use SELECT) and by using another feature (here MARKING) to, somewhat redundantly in the presence of valence features, directly describe saturation. This eliminates the features MOD, SPEC and SPR and turns out to require fewer syntactic schemata.

A head-functor schema then comes into play, which, as outlined in Figure 1, identifies the element in the functor's SELECT feature with the head daughter. While the mother node's valence and head features are shared with the head element, as expected, its MARKING value is contributed by the functor, via functor's MARK value.[5]

A type hierarchy for the possible values that features MARKING and MARK can take are then used to enforce the subtleties of the NP internal syntactic structure. For the sake of illustration, a type hierarchy for MARKING values is presented in Figure 2, and the following paragraphs describe the necessary constraints where they are employed to model a very simplified NP structure assumed in the remaining Sections.[6]

---

[3]Ginzburg and Sag (2001) make heavy use of them, also in the analyses of constructions related to ellipsis, like sluicing, and Sag (2000) employs a syntactic rule to handle VP ellipsis that in some cases may be unary.

[4]This is in line with similar proposals in transformational grammar, including the influential Kayne (1994), which assumes perhaps the more far reaching motto that "specifiers are an instance of adjunction".

[5]Despite the name of MARKING we are using for this feature, functors are not to be confused with the markers of Pollard and Sag (1994).

[6]This hierarchy will be used for all languages under consideration. Inevitably, if one wants to cover phenomena that are not treated in this paper, there will be language-related differences. To give an example, in Portuguese, possessives co-occur with articles: $o_{Article}\ teu_{Possessive}\ carro_{Noun}$

$$\begin{bmatrix} \textit{head-functor-phr} \\ \text{SS|LOC|CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{VAL} & \boxed{2} \\ \text{MARKING} & \boxed{3} \end{bmatrix} \\ \text{HEAD-DTR|SS} \boxed{4} \begin{bmatrix} \text{LOC|CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{VAL} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{FUNCTOR-DTR|SS|LOC|CAT|HEAD} \begin{bmatrix} \text{MARK} & \boxed{3} \\ \text{SELECT} & \boxed{4} \end{bmatrix} \end{bmatrix}$$
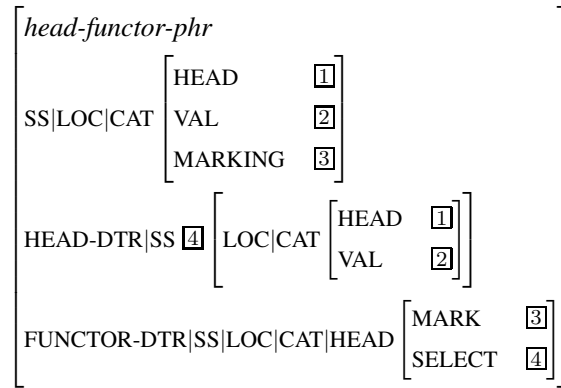
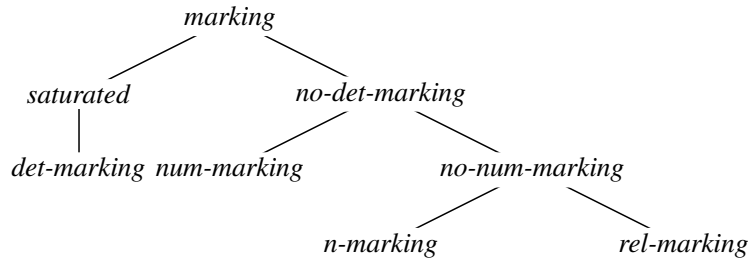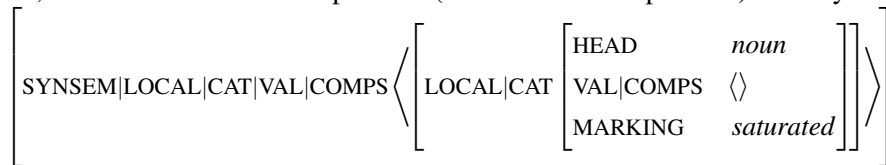Figure 1: Outline of head-functor schema

Figure 2: Type hierarchy of marking

With this setup, items that select for NPs constrain them to have a MARKING with value *saturated* (instead of requiring their SPR feature to be empty). For instance, an item with an NP complement (and no other complement) will say:

$$\begin{bmatrix} \text{SYNSEM|LOCAL|CAT|VAL|COMPS} \left\langle \begin{bmatrix} \text{LOCAL|CAT} \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{VAL|COMPS} & \langle \rangle \\ \text{MARKING} & \textit{saturated} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Nouns come in the lexicon with [MARKING *n-marking*]. Therefore, they need to combine with a functor with a MARK value unifiable with *saturated* in order for the resulting constituent to be able to occur in NP contexts.

Determiners select a constituent with a value of MARKING incompatible with the value of their MARK feature, so that they do not iterate:

$$\begin{bmatrix} \text{SYNSEM|LOCAL|CAT|HEAD} \begin{bmatrix} \textit{determiner} \\ \text{SELECT|LOCAL|CAT|MARKING} & \textit{no-det-marking} \\ \text{MARK} & \textit{det-marking} \end{bmatrix} \end{bmatrix}$$

Numerals are similar but less peripheral:

$$\left[\text{SYNSEM|LOCAL|CAT|HEAD} \begin{array}{l} \textit{numeral} \\ \text{SELECT|LOCAL|CAT|MARKING} \quad \textit{no-num-marking} \\ \text{MARK} \qquad\qquad\qquad\qquad\quad \textit{num-marking} \end{array} \right]$$

APs and PPs select for a constituent with [MARKING *n-marking*] and produce a node with the same level of saturation:

$$\left[\text{SYNSEM|LOCAL|CAT|HEAD} \begin{array}{l} \textit{adjective} \lor \textit{preposition} \\ \text{SELECT|LOCAL|CAT|MARKING} \quad \textit{n-marking} \\ \text{MARK} \qquad\qquad\qquad\qquad\quad \textit{n-marking} \end{array} \right]$$

As a consequence, they are allowed to recur and, when following the noun, to be interspersed (i.e. both Noun-AP-PP and Noun-PP-AP sequences are grammatical in languages that allow postnominal adjectives).[7]

Relative clauses should be allowed to iterate, but they are more peripheral than APs and PPs inside an NP:

$$\left[\text{SYNSEM|LOCAL|CAT|HEAD} \begin{array}{l} \text{SELECT|LOCAL|CAT|MARKING} \quad \textit{no-num-marking} \\ \text{MARK} \qquad\qquad\qquad\qquad\quad \textit{rel-marking} \end{array} \right]$$

Appropriate nesting is thus enforced with the help of MARKING values: for instance, determiners are more peripheral than numerals.

Marking values do not constrain relative word order between the daughters of head-functor phrases, which must be enforced separately.[8]

The use of a unary schema can also be assumed to produce bare NPs when appropriate. Instead of discharging the SPR feature of nouns, that rule takes as daughter a noun-headed constituent with [MARKING *no-det-marking*] and produces a node with [MARKING *det-marking*].

3.1 Example

An example parse for the NP *these two cars* is presented in Figure 3.

---

[7]If a language has both prenominal and postnominal adjectives, or prenominal adjectives and postnominal PPs, potentially spurious attachment ambiguities will be produced for a sequence AP-Noun-PP/AP: [AP [Noun PP/AP]] and [[AP Noun] PP/AP]. It is straightforward to complicate the type hierarchy of *marking* to control this, too. If one wants to keep the structure [AP [Noun PP/AP]] and rule out [[AP Noun] PP/AP], the type *n-marking* can be split into two subtypes *pre-n-marking* and *post-n-marking*, and prenominal adjectives can be specified to have the constraint [MARK *pre-n-marking*] and select for nominal projections with [MARKING *n-marking*], with prepositions and postnominal adjectives selecting for sister nodes with [MARKING *post-n-marking*] and bearing an identical value for their MARK attribute. Nouns would then come in the lexicon with [MARKING *post-n-marking*]. We will ignore this complication in the remainder of this text.

[8]For instance, by having two subtypes of the functor-head schema with different precedence relations between head and functor and controlling rule application by some dedicated feature in functors.
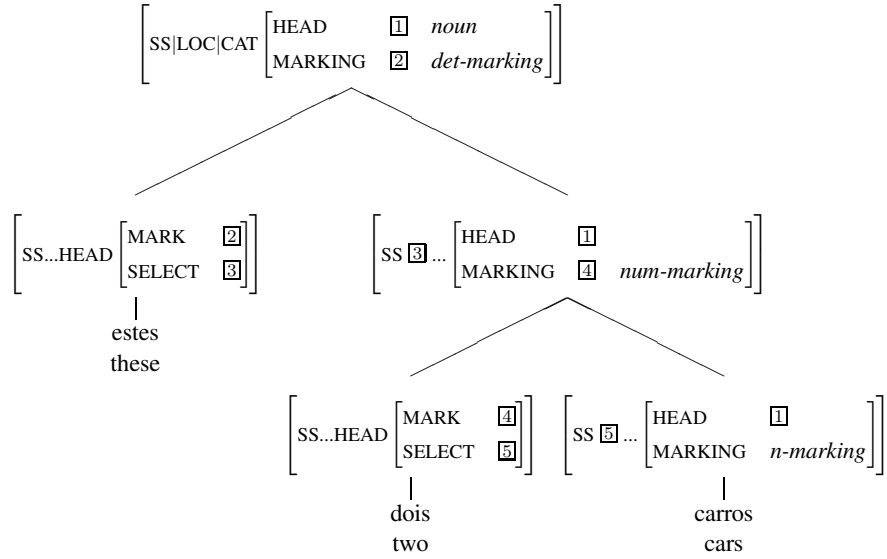
$$\begin{bmatrix} \text{SS|LOC|CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} & \textit{noun} \\ \text{MARKING} & \boxed{2} & \textit{det-marking} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{SS...HEAD} \begin{bmatrix} \text{MARK} & \boxed{2} \\ \text{SELECT} & \boxed{3} \end{bmatrix} \end{bmatrix} \qquad \begin{bmatrix} \text{SS} \boxed{3} ... \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{MARKING} & \boxed{4} & \textit{num-marking} \end{bmatrix} \end{bmatrix}$$

estes
these

$$\begin{bmatrix} \text{SS...HEAD} \begin{bmatrix} \text{MARK} & \boxed{4} \\ \text{SELECT} & \boxed{5} \end{bmatrix} \end{bmatrix} \qquad \begin{bmatrix} \text{SS} \boxed{5} ... \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{MARKING} & \textit{n-marking} \end{bmatrix} \end{bmatrix}$$

dois   carros
two   cars

Figure 3: Parse tree for *estes/these dois/two carros/cars*.

4 Analysis

In the approach sketched in the previous Section, both for specifiers and adjuncts, the information about their head can be found in a single place (the SELECT feature), and the same holds for the information on the nature of the constituents they yield when they are attached to their head (under the MARK feature). This account of NPs in general brings two important advantages: (1) specifiers and modifiers receive a uniform treatment; (2) since all the syntactic properties of the constituent resulting from the attachment of a functor with its head are present in the functor, they will be known if the head is missing. Therefore, a single schema for noun ellipsis can be implemented for both specifiers and adjuncts ensuring syntactic structures that replicate the ones obtained when the nominal head is not missing.

Against this background, elliptical NPs without complements can be easily accounted for with the help of a syntactic schema *n-ellipsis-functor*, which is a straightforward unary version of the schema in Figure 1 for NPs but without the HEAD-DTR. Some properties of this schema are:

- the MARKING value of the mother node is given by its functor's MARK value;

- the SYNSEM of the mother node is partly shared with the SYNSEM of the functor's SELECT value: it is shared at least for the features HEAD and VAL.[9] As for the remainder features, note that, on the one hand, the SYNSEM|LOCAL|

---

[9]These are the same features that are shared between the mother and the head-daughter in a head-functor phrase, and functors must be able to fully specify the level of saturation of the head they attach to.
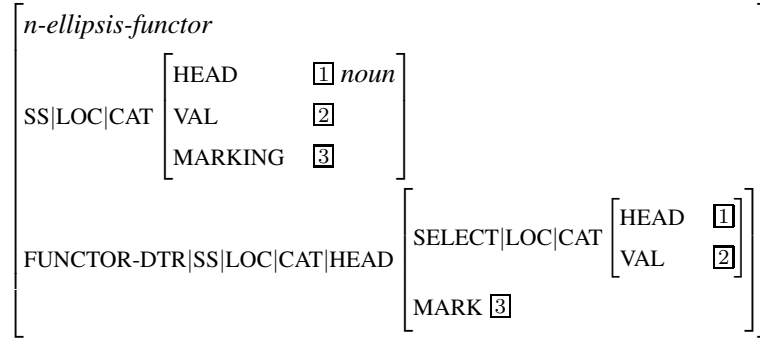
$$\begin{bmatrix} \textit{n-ellipsis-functor} \\ \text{SS|LOC|CAT} \begin{bmatrix} \text{HEAD} & \boxed{1}\ \textit{noun} \\ \text{VAL} & \boxed{2} \\ \text{MARKING} & \boxed{3} \end{bmatrix} \\ \text{FUNCTOR-DTR|SS|LOC|CAT|HEAD} \begin{bmatrix} \text{SELECT|LOC|CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{VAL} & \boxed{2} \end{bmatrix} \\ \text{MARK}\ \boxed{3} \end{bmatrix} \end{bmatrix}$$

Figure 4: Outline of the noun ellipsis schema.

CONT|RELS of the mother node must be the union of the functor's RELS with a multi-set of relations corresponding to the denotation of the missing noun; on the other hand, the MARKING values (i.e. the MARKING feature of the mother node and the MARKING feature of the synsem in the SELECT attribute of the functor) may be incompatible and should not be shared at all;

- the HEAD of the mother is constrained to be a noun (functors not selecting nouns via the SELECT feature will thus not be part of an elliptical NP), and its COMPS should be inherited from the antecedent.

Hence, given an elliptical NP, this schema will directly apply to the functor with the most specific marking type. The other functors will be combined as expected, following the usual schemata in place also for non elliptical NPs.

Figure 4 depicts the syntactic constraints associated with the noun ellipsis schema. The semantic properties of this construction are discussed in Section 4.3.

4.1 Example

We present the parse tree for the NP *estes dois/ these two* in Figure 5. The numeral *dois/two* feeds the *n-ellipsis-functor* rule and yields a node with [HEAD *noun*] and [MARKING *num-marking*]. The determiner attaches as expected, via some head-final version of the head-functor schema in Figure 1, giving rise to a node with [MARKING *det-marking*], a full (saturated) NP. The resulting structure is completely parallel to the one of an NP like *estes dois carros* (*these two cars*), except for the missing N node and the branch connecting it (Figure 3).

The NP *estes/these* will also be generated (with *estes/these* feeding the ellipsis rule and producing a projection headed by a noun with a saturated MARKING value, i.e. an NP), but then the numeral cannot attach for the very reasons that prevent numerals from modifying full NPs: conflicting values of MARKING and other constraints that also block numerals from attaching to the right of constituents in Portuguese and English.

In general, NPs with elided nouns are derived by an application of the ellipsis rule to the most embedded constituent as defined by the marking hierarchy. The

$$
\left[\begin{array}{l}
\text{SS|LOC|CAT} \left[\begin{array}{ll}
\text{HEAD} & \boxed{1}\ \textit{noun} \\
\text{VAL} & \boxed{2} \\
\text{MARKING} & \boxed{5}\ \textit{det-marking}
\end{array}\right]
\end{array}\right]
$$

$$
\left[\begin{array}{l}
\text{SS|LOC|CAT|HEAD} \left[\begin{array}{ll}
\text{SELECT} & \boxed{4} \\
\text{MARK} & \boxed{5}
\end{array}\right]
\end{array}\right]
$$

estes
these

$$
\left[\begin{array}{l}
\text{SS}\ \boxed{4}\ \text{LOC|CAT} \left[\begin{array}{ll}
\text{HEAD} & \boxed{1}\ \textit{noun} \\
\text{VAL} & \boxed{2} \\
\text{MARKING} & \boxed{3}\ \textit{num-marking}
\end{array}\right]
\end{array}\right]
$$

$$
\left[\begin{array}{l}
\text{SS|LOC|CAT|HEAD} \left[\begin{array}{l}
\text{SELECT...} \left[\begin{array}{ll}
\text{HEAD} & \boxed{1} \\
\text{VAL} & \boxed{2}
\end{array}\right] \\
\text{MARK}\ \boxed{3}
\end{array}\right]
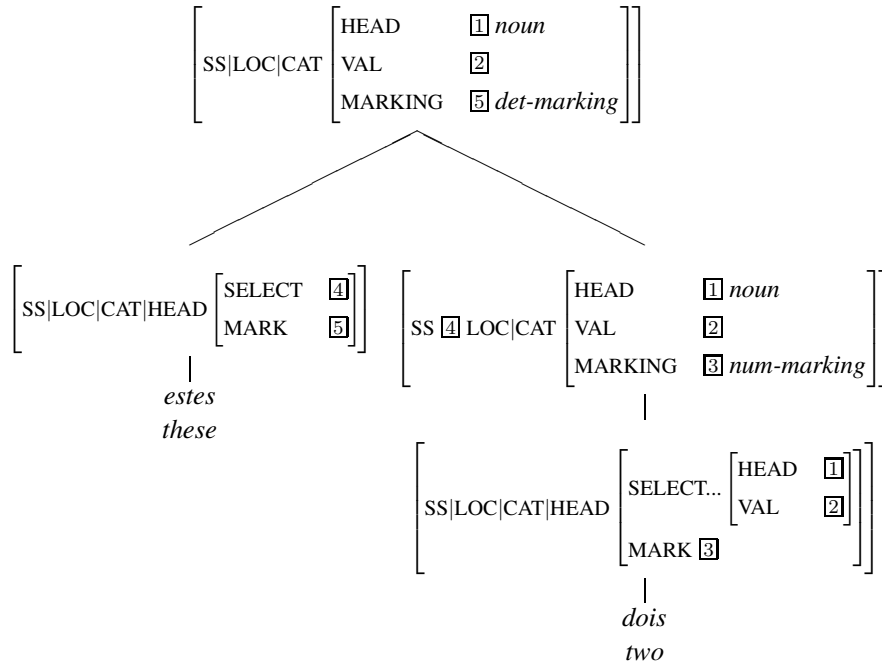\end{array}\right]
$$

dois
two

Figure 5: Parse tree for the NP [*estes/these dois/two* -].

other functors that are present combine as expected. In (9) we show the structures produced by the present analysis for some of the English ((9a) and (9b)), Spanish (9c) and German (9d) examples in Section 1:

(9) a. [$_{det-marking}$ [$_{D}$ some]]

b. [$_{det-marking}$ [$_{D}$ her] [$_{num-marking}$ [$_{Num}$ two]]]

c. [$_{det-marking}$ [$_{D}$ la] [$_{n-marking}$ [$_{A}$ verde]]]

d. [$_{det-marking}$ [$_{n-marking}$ [$_{A}$ junge] [$_{n-marking}$ [$_{PP}$ mit Mütze]]]]

4.2 Antecedent Resolution

The relation between an NP with an elided noun and its antecedent has been reported in the literature to have properties in common with the kind of binding ruled by Principle B (Hankamer and Sag (1976), Lobeck (1995) among others; the following examples are theirs). In fact, the antecedent can be given pragmatically, as in (10a), or be in a different sentence (10b).

(10) a. *At a food vendor's*: I'll take [two -].

b. - John caught a big fish.
   - Yes, but [Mary's -] was bigger.

The way to determine the antecedent may thus be dependent on how anaphoric binding is analyzed in general and will not be discussed here. But it is worth noting that, whereas in binding there is an anaphoric relation between NPs, here there is a semantic dependency relation between predicators (sentence (3a) illustrates this point clearly). We will continue using the expression *one anaphora* in this text though.

4.3 Semantics

Bearing in mind the considerations in the previous paragraphs, we illustrate the composition of semantics for these phrases, ignoring for now the way the antecedent is to be recovered. Minimal Recursion Semantics (Copestake et al. (2005)) is used to this end.[10] Figure 6 shows the semantic constraints on *n-ellipsis-functor*, and its main properties are the following:

- the SYNSEM|LOCAL|CONT|RELS of the mother node is the union of the functor's RELS with a multi-set with a nominal object in it (a *noun-ellipsis* relation);

- since no handle constraints should be associated with the missing noun, the HCONS feature of the mother node is simply the HCONS of the daughter;

- the INDEX of a nominal projection is the INDEX of the head noun, which is structure-shared with the ARG0 of the noun's relation in the lexical entry for the noun. In the absence of this lexical unit, this unification must be performed here, by directly identifying the INDEX of the mother node with the ARG0 of the *noun-ellipsis* relation;

- the functor must be allowed to see the LTOP and the INDEX of the node it selects because they can be arguments of the relation or relations the functor contributes to the semantics. Since a noun would equate its LTOP with the LBL feature of its relation and its INDEX with the ARG0 feature there, these are unified with the LTOP and INDEX under the SELECT attribute of the functor;

- to simplify our presentation, we ignore Kasper's problem (Copestake et al. (2005)) in this analysis and (1) unify the LTOP of the mother node with the LTOP of the daughter, and (2) assume in what follows that, in the lexicon, intersective modifiers identify their LTOP with the LTOP of what they select. It should be clear that the present proposal is compatible with any of the two known solutions to that problem (i.e. multiplying syntactic rules or multiplying the features used for the composition of semantics; see the cited paper).

---

[10]For conciseness reasons, we omit feature HOOK in our presentation.

$$
\begin{bmatrix}
\textit{n-ellipsis-functor} \\[4pt]
\text{SS|LOC|CONT}
\begin{bmatrix}
\text{LTOP} & \boxed{1} \\
\text{INDEX} & \boxed{3} \\[6pt]
\text{RELS} & \boxed{A} \cup \left\{
\begin{bmatrix}
\textit{noun-ellipsis\_rel} \\
\text{LBL} & \boxed{2} \\
\text{ARG0} & \boxed{3}
\end{bmatrix}
\right\} \\[6pt]
\text{HCONS} & \boxed{B}
\end{bmatrix} \\[10pt]
\text{FUNCTOR-DTR|SS|LOC}
\begin{bmatrix}
\text{CAT|HEAD|SELECT|LOC|CONT}
\begin{bmatrix}
\text{LTOP} & \boxed{2} \\
\text{INDEX} & \boxed{3}
\end{bmatrix} \\[6pt]
\text{CONT}
\begin{bmatrix}
\text{LTOP} & \boxed{1} \\
\text{RELS} & \boxed{A} \\
\text{HCONS} & \boxed{B}
\end{bmatrix}
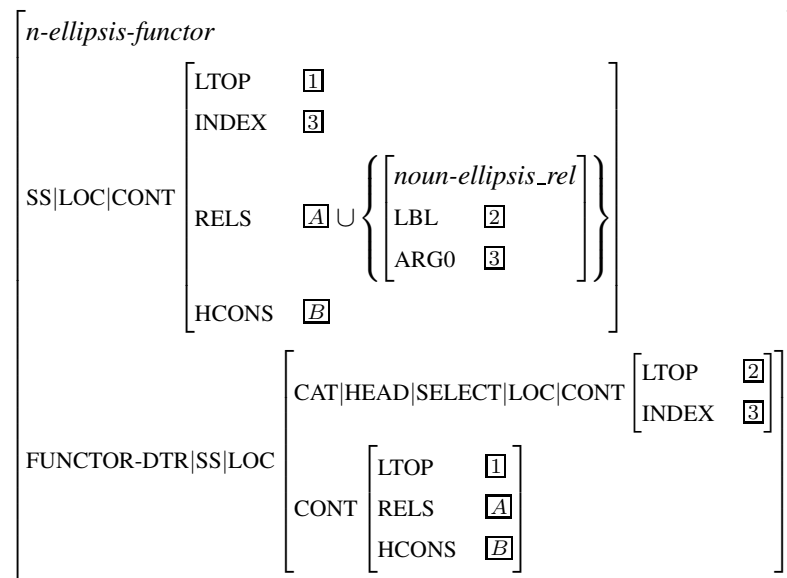\end{bmatrix}
\end{bmatrix}
$$

Figure 6: Semantic constraints of the noun ellipsis schema.

4.3.1 Example

We present an example parse for the NP *some - in Varna*, decorated with LTOP and INDEX features, in Figure 7. In that figure, it is assumed that features SS|LOC|CAT| HEAD|SELECT|LOC|CONT|LTOP and SS|LOC|CONT|LTOP are unified in the lexical entry for the preposition. The resulting MRS is presented in Figure 8.

4.4 Structural Parallelism

Maintaining structural parallelism between NPs with expressed nouns and NPs with missing nouns has several advantages. For instance, if we assume that relative clauses attach lower than determiners in expressed noun NPs, as we have been, in the present analysis an NP like *some that arrived* will get the parallel structure [[*some* D] [[*that arrived* RC] $\overline{\text{N}}$] NP]. This can be important for semantic reasons, since restrictive relative clauses contribute relations in the restrictor of the quantifier of the NP they are in.

An alternative analysis where *some* would be, say, a pronoun and restrictive modifiers must attach higher would introduce asymmetries in the way semantics is built. For example, in the analysis of Beavers (2003), a unary syntactic rule is put in place for noun ellipsis that takes a determiner as input, and noun-headed projections are considered complements of determiners (the DP hypothesis is followed). This analysis has obvious problems attaching a relative clause (or a numeral as in *these three*, etc.) to an unexpressed complement of a determiner and will thus be forced to have these elements attached to full DPs when no noun is present. In the analysis presented here, determiners, with [MARK *det-marking*], produce full
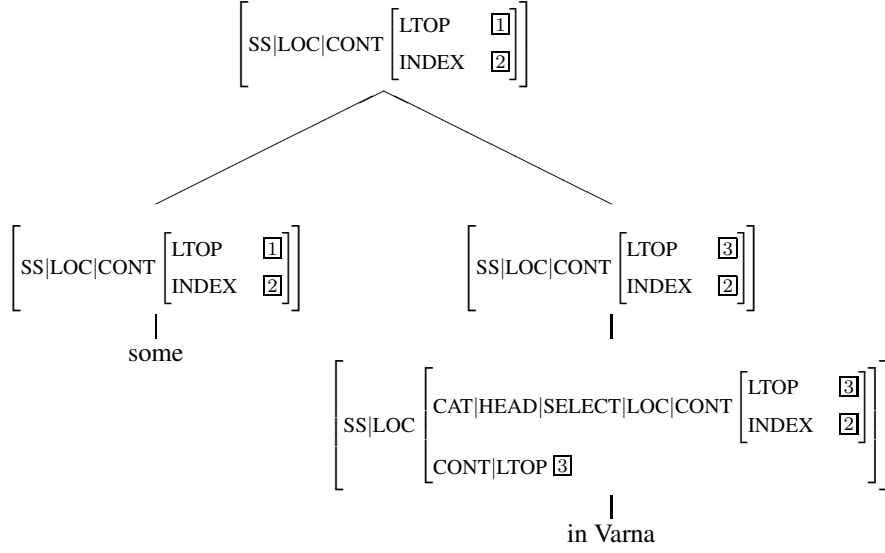
$$
\left[\text{SS|LOC|CONT} \begin{bmatrix} \text{LTOP} & \boxed{1} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \right]
$$

$$
\left[\text{SS|LOC|CONT} \begin{bmatrix} \text{LTOP} & \boxed{1} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \right]
$$

some

$$
\left[\text{SS|LOC|CONT} \begin{bmatrix} \text{LTOP} & \boxed{3} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \right]
$$

$$
\left[\text{SS|LOC} \begin{bmatrix} \text{CAT|HEAD|SELECT|LOC|CONT} \begin{bmatrix} \text{LTOP} & \boxed{3} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \\ \text{CONT|LTOP } \boxed{3} \end{bmatrix} \right]
$$

in Varna

Figure 7: Parse for the example NP *some - in Varna*.

$$
\begin{bmatrix}
\text{LTOP} & \boxed{h7}\, h \\
\text{INDEX} & \boxed{x6}\, x \\
\text{RELS} & \left\langle
\begin{bmatrix} \textit{some\_rel} \\ \text{LBL} & \boxed{h7} \\ \text{ARG0} & \boxed{x6} \\ \text{RSTR} & \boxed{h9}\,h \\ \text{BODY} & \boxed{h8}\,h \end{bmatrix},
\begin{bmatrix} \textit{noun-ellipsis\_rel} \\ \text{LBL} & \boxed{h10}\,h \\ \text{ARG0} & \boxed{x6} \end{bmatrix},
\begin{bmatrix} \textit{in\_rel} \\ \text{LBL} & \boxed{h10} \\ \text{ARG1} & \boxed{x6} \\ \text{ARG2} & \boxed{x11}\,x \end{bmatrix},
\right. \\
& \left.
\begin{bmatrix} \textit{proper\_rel} \\ \text{LBL} & \boxed{h13}\,h \\ \text{ARG0} & \boxed{x11} \\ \text{RSTR} & \boxed{h15}\,h \\ \text{BODY} & \boxed{h14}\,h \end{bmatrix},
\begin{bmatrix} \textit{named\_rel} \\ \text{LBL} & \boxed{h16}\,h \\ \text{ARG0} & \boxed{x11} \\ \text{CARG} & \textit{Varna} \end{bmatrix}
\right\rangle \\
\text{HCONS} & \left\langle
\begin{bmatrix} \textit{qeq} \\ \text{HARG} & \boxed{h9} \\ \text{LARG} & \boxed{h10} \end{bmatrix},
\begin{bmatrix} \textit{qeq} \\ \text{HARG} & \boxed{h15} \\ \text{LARG} & \boxed{h16} \end{bmatrix}
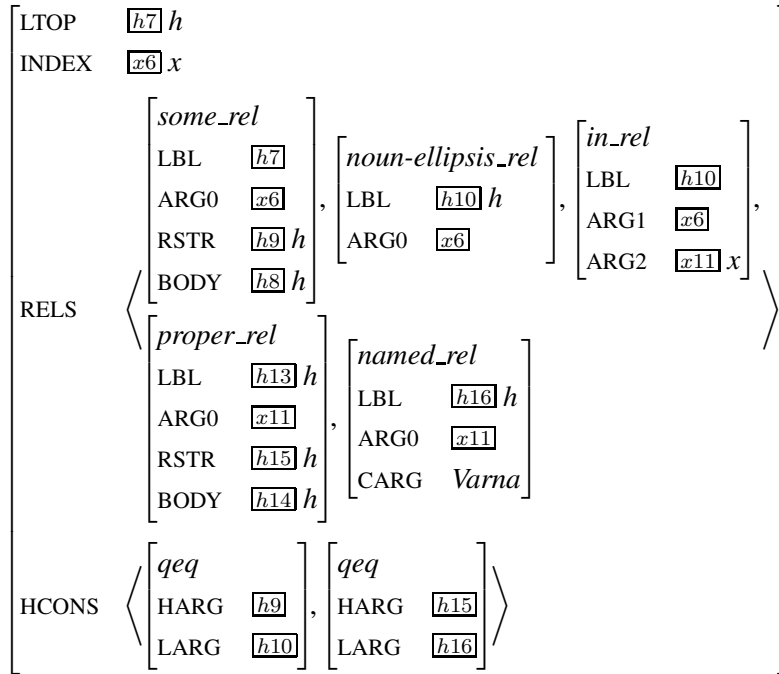\right\rangle
\end{bmatrix}
$$

Figure 8: MRS for the NP *some - in Varna*.

NPs when they undergo the ellipsis rule, but if a relative clause is present, it is this element that feeds *n-ellipsis-functor* and the determiner attaches higher.

To put it more explicitly, a restrictive relative clause will behave like the PP in Figure 7, identifying its LTOP with the LTOP of the constituent it selects. In an analysis where determiners of elided nouns are treated as pronouns, a relative clause would have to attach to a full NP. In that case, the LTOP of its sister would have a different value, as can be seen in that Figure, and extra features would be required to fix the problem. Allowing for NPs like [*these two* -] and simultaneously blocking numerals from attaching to the right of NPs headed by overt nouns would be even more cumbersome.

4.5 Complements of Null Nouns

The COMPS feature of the node produced by the *n-ellipsis-functor* rule should be the same as the COMPS feature of the antecedent noun: in languages where noun modifiers can intervene between the noun and its complements, head-complement schemata can apply higher, for instance to derive Portuguese examples like:[11]

(11) o   filho mais velho do   Rui e   o   [$_{\overline{N}}$ [$_N$ - mais novo] [da   Ana]]
      the son   most old   of the Rui and the   most young   of the Ana
      Rui's eldest son and Ana's youngest one

However, since the deepest constituents of elliptical NPs with complements might not be functors (i.e. might be the complements themselves), a further unary schema, *n-ellipsis-comp* is required. For example, we want to assign to the NP *o da Ana* the structure presented in (12), but the most embedded PP is not a functor.

(12) o   filho do   Rui e   [$_{NP}$ o   [$_{\overline{N}}$ - [$_{PP}$ da   Ana]]]
      the son   of the Rui and   the   of the Ana
      Rui's son and Ana's

This second schema simply turns a PP that can be a noun complement into a nominal projection. The remaining functors, if present, will attach upwards as expected. This schema should of course ensure that the complement PP is compatible with the selectional properties inherited from the antecedent of the elliptical NP.

Some of its key properties are common to the *n-ellipsis-functor* schema above. The resulting constituent:

- has a HEAD of type *noun* and a MARKING of type *n-marking*;

- since the antecedent can have more than one complement, its COMPS value is the tail of the COMPS value of the antecedent of the elliptical NP;

---

[11]This is why in *n-ellipsis-functor* the VAL of the mother node is shared with the VAL in the functor's SELECT feature: functors that must attach only after the complements are projected — e.g. all adnominals in English and relative clauses in Portuguese — can specify that they attach to a projection with empty COMPS, with the desired effect that they are blocked from appearing in contexts like (11).
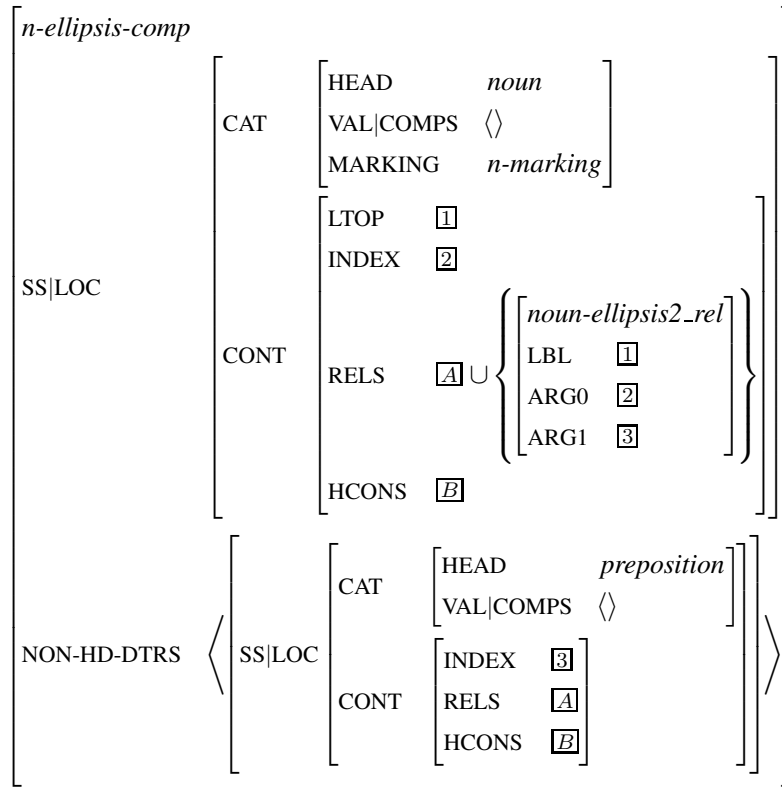
$$
\begin{bmatrix}
\textit{n-ellipsis-comp} \\[2pt]
\text{SS|LOC}\begin{bmatrix}
\text{CAT}\begin{bmatrix}
\text{HEAD} & \textit{noun} \\
\text{VAL|COMPS} & \langle\,\rangle \\
\text{MARKING} & \textit{n-marking}
\end{bmatrix} \\[4pt]
\text{CONT}\begin{bmatrix}
\text{LTOP} & \boxed{1} \\
\text{INDEX} & \boxed{2} \\
\text{RELS} & \boxed{A} \cup \left\{ \begin{bmatrix}
\textit{noun-ellipsis2\_rel} \\
\text{LBL} & \boxed{1} \\
\text{ARG0} & \boxed{2} \\
\text{ARG1} & \boxed{3}
\end{bmatrix} \right\} \\[4pt]
\text{HCONS} & \boxed{B}
\end{bmatrix}
\end{bmatrix} \\[6pt]
\text{NON-HD-DTRS}\ \left\langle \begin{bmatrix}
\text{SS|LOC}\begin{bmatrix}
\text{CAT}\begin{bmatrix}
\text{HEAD} & \textit{preposition} \\
\text{VAL|COMPS} & \langle\,\rangle
\end{bmatrix} \\
\text{CONT}\begin{bmatrix}
\text{INDEX} & \boxed{3} \\
\text{RELS} & \boxed{A} \\
\text{HCONS} & \boxed{B}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix} \right\rangle
\end{bmatrix}
$$

Figure 9: Additional noun ellipsis schema for PPs as noun complements.

- its RELS and HCONS values are also mainly contributed by the daughter, with the semantics of the missing noun added;

- the LTOP and INDEX features of the mother node correspond to the features LBL and ARG0 of the added relation, as in nouns.

A few technical issues arise if the antecedent is not available: (1) it is not possible to constrain the value of the COMPS in the mother node appropriately; (2) it is not possible to know the arity of the relation for the missing noun; and (3) it is not possible to know how the arguments of that relation are instantiated (e.g. the second argument of the nominal relation is given by the INDEX feature of the daughter node if it is a PP headed by a non-predicational preposition, but by its LTOP feature if it is a CP). In general, selectional properties of the missing noun are not known but they are required to constrain the daughter appropriately. The antecedent is thus crucial to resolve these values. Figure 9 depicts this schema, limited only to PP complements of nouns selecting for a single complement. If the antecedent cannot be known, additional schemata may be necessary for other kinds of complements.

Note that this second schema may not be required for every language. Lobeck (1995) points out that in English, elliptical NPs cannot contain noun complements.

Therefore, in some languages *n-ellipsis-functor* may suffice to account for all data concerning noun ellipsis.[12]

4.6 Co-Occurrence Restrictions

In order to control many co-occurrence restrictions in elliptical NPs, the *Left Periphery (Empty)* (LP(E)) attribute is used in some analyses (Nerbonne and Mullen (2000), Netter (1996)).[13] Such an approach is compatible with the present proposal and can be kept in use for the same purpose.

In (Nerbonne and Mullen, 2000), null nouns have the value *null* for their LP feature, the anaphoric *one* shows the value *one* and all other lexical items are constrained to be [LP *full*], a value that is percolated from the leftmost daughter in phrases. In the following paragraphs we show how this behavior can be mimicked when one uses unary rules instead of the null noun.

In order to make use of LP in this approach, each ellipsis rule can be split into two rules: one that corresponds to the case where the missing noun precedes the functor or complement, the other corresponding to the inverse situation. Note that not all combinations will be necessary for every language, as in languages with fixed head-complement word order only one *n-ellipsis-comp* rule will be needed. Head-initial ellipsis rules must then be specified to be [LP *empty*], the others structure-share the LP feature in the mother node with the LP attribute of the functor daughter (which will be *full*). Control on which functors feed each ellipsis rule can be done in the same way as control on linear precedence between a functor and an expressed head.[14]

An additional constraint must be added to *n-ellipsis-functor* rules — their functor daughter selects an element with [LP *empty*], because the LP(E) analyses rely on functors being able to see the LP value of what they select, and we want to mimic the effect of them attaching to a null constituent:

$$\left[\text{FUNCTOR-DTR|SS|LOC|CAT|HEAD|SELECT|LP } \textit{empty} \right]$$

The LP constraints on lexical items and remaining phrases are as in the original proposals.

Sections 4.6.1, 4.6.2, 4.6.3 and 4.6.4 make use of the LP machinery to tackle the co-occurrence restrictions of elided nouns enumerated in Section 1.

4.6.1 NP Initial Ellipsis

In some languages, like German, where NP initial ellipsis is ruled out (ex. (3)), LP has been used to prevent PPs from being NP initial. That analysis can be used here. Bare-NPs can be produced via a unary syntactic rule that adds quantifier semantics

---

[12]Additionally, the COMPS of the mother node in *n-ellipsis-functor* may have to be constrained to be empty in these languages. In English, it is actually not required, since nominal functors select synsems with empty COMPS anyway, and *n-ellipsis-functor* is sensitive to this information (Figure 4).

[13]In Nerbonne and Mullen (2000) LP takes the values *empty*, *full* or *one*. We will also use these.

[14]We are assuming a setup like the one described in footnote 8.

and produces a node with [MARKING *det-marking*]. In these languages the bare-NP rule must also constrain its daughter to be [LP *full*].

4.6.2 *One* Anaphora vs. Missing Nouns

The LP feature used in (Nerbonne and Mullen, 2000) to account for the specific distribution of NPs with *one* and NPs with a missing head noun in English are compatible with the present proposal, and will keep ensuring the same results in this respect.[15]

4.6.3 Definite Articles

In some languages, like English or Portuguese, some specifiers like the definite article cannot alone form an elliptical NP (ex. (4)).

The LP feature has been used also to promote this blocking effect. Again, this analysis can also be incorporated here: the definite articles can select an element with [LP ¬*empty*]. More on this will be said in Section 5.

4.6.4 Prenominal Adjectives of Romance Languages

In some languages, like Spanish or Portuguese, predominantly with post-head adjectives, pre-head ones cannot feed the ellipsis rules (ex. (5)).

The LP analysis can be extended to accommodate these data straightforwardly. All that is needed is that these adjectives select a constituent that is [LP ¬*empty*].[16]

It should be noted that it is not possible to test how local this blocking effect is, i.e. whether these adjectives are really sensitive to edge constraints or to the absence of the head noun, since the only material that can intervene between a prenominal adjective and the noun is another prenominal adjective.

5   Problems with the LP Analysis

However successful it may be for most of the issues tackled above, the Left Periphery analysis of Nerbonne and Mullen (2000) makes wrong predictions regarding the distribution of the English and the Portuguese definite articles. Below, we identify two problems that seem to be manifestations of the same underlying issue. But note that this is orthogonal to whether null categories are used or not.

The first problem is related to the fact that Portuguese simultaneously (1) lacks *one* anaphora, (2) does not allow definite articles to make up NPs alone, but (3) allows them to co-occur with postnominal material in elliptical NPs.

---

[15]For instance, with the setup presented above, it can be maintained that the English "many" selects a constituent with [LP ¬*one*] and "none" selects one with [LP *empty*]. See Nerbonne and Mullen (2000) for several examples.

[16]Recall that *n-ellipsis-functor* rules constrain its functor daughter to select an element with [LP *empty*].

Whereas the constraint presented in Section 4.6.3 is appropriate for English, it may not be for other languages: although NPs that consist of a single article are blocked because of that constraint, other NPs are wrongly so as well. Consider the Portuguese example:

(13) Os homens de chapéu encontraram [$_{NP}$ os [$_{\overline{N}}$ - de boné]].
     the men   of hat   met   the   of cap
     The men in hats met the ones in caps.

The constituent selected by the determiner (labeled $\overline{N}$ in the example) actually has the property [LP *empty*] under all analyses using the LP feature, because PPs attach to the right of nouns. The same also happens in all NPs where a definite article immediately precedes a postnominal functor (be it a PP, an AP or a relative clause) or a nominal complement. The LP analysis thus wrongly rejects these NPs.

Note that the distribution of the Portuguese definite article is independent of whether the noun is realized or not, because of NPs like [*os dois (carros)*] ([*the two (cars)*]), and it is also independent of edge features, since the article can attach to [LP *full*] constituents and to [LP *empty*] ones.

The second problem arising from the usage of edge features concerns the English article: we cannot simultaneously accept an NP like *the one in caps* and block an example like:[17]

(14) * A single protester attended the rally because [the one] apparently felt it was important.

If we consider these two problems together and use the empty categories metaphor for ease of exposition, it seems that the definite article in these languages must attach to a constituent which contains more material than just the null noun or the anaphoric *one*. In the case of English (but not of Portuguese) the sister of the article is also required to be [LP ¬ *empty*], as presented in Section 4.6.3.

It is not clear what sets the two constructions just spotted above apart from the rest. We think that it is not a phonological or semantic issue. It cannot be a phonological question, because there is nothing special with the item *one* compared to nouns in this respect, and in fact the English NP *the one* is attested when the item *one* is not the anaphoric *one*.

In turn, if it were a semantic effect, it would be a surprise that some languages may allow it. If one accepts that the form *der* in example (15) is in fact an article and not a demonstrative, German is one such language.

(15) Wir haben einen Film   gesehen aber [der -] war langweilig.
     we   have a   movie watched but   the   was boring
     We watched a movie but it was boring.

---

[17]If the English definite article attaches to a constituent that is [LP ¬ *empty*], then *the one* is accepted; if its sister is constrained to be [LP *full*], then *the one in caps* is rejected.

It is thus possible that this constraint is a purely formal requirement. We abstain however from defending the trivial technical solution of distinguishing these two constructions from the rest via some feature that shows one value in one group and an incompatible one in the other, given its stipulative nature.

6 Related Issues and Open Questions

The analysis of noun ellipsis presented here can be extended also to related constructions. For instance, partitive constructions of the form *some of them* can be analyzed as elliptical and accounted for in a similar and direct fashion.

The present analysis also straightforwardly covers the syntax of missing-N generics (e.g. *the poor*). Their semantics, however, should be different from the semantics of elliptical NPs, since they do not have an antecedent, typically denote humans and carry kind readings.

In this connection, it is worth noting that the difference between the two constructions (NP ellipsis and missing-N generics) also involves lexical idiosyncrasies. For example, the Portuguese NP in (16a) is ambiguous between the noun ellipsis and the missing-N generic reading, as its two English correlates indicate. The English correlate with *one* anaphora corresponds to the ellipsis reading, and the English correlate with a missing noun corresponds to the missing-N generic reading.

(16) a. [os pobres -]
           the poor
        the poor   (missing-N generic reading)
        the poor ones   (noun ellipsis reading)

      b. [os dois -]
           the two
        the two   (noun ellipsis reading)

The NP in (16b) lacks the missing-N generic reading, and, accordingly, only has one English correlate. But in this case, English surprisingly uses the missing noun strategy, although one would expect noun ellipsis readings to correspond to *one* anaphora here, too. Future research may use such considerations to shed light on the distribution of *one* (only used in the languages that have them when there could otherwise be ambiguity between the two constructions?), which is accounted for by the LP analyses in a completely stipulative way.

Another issue that is left open is the status of personal pronouns. The point here is whether personal pronouns are fully saturated NPs or rather determiners occurring in NPs missing the noun. Phrases like the English *you two* or *we students* might suggest the latter answer, but personal pronouns fail to systematically show the ability to be restrictively modified.

7 Conclusion

In this paper, we outlined an analysis of noun ellipsis that builds on the selectional information lexically available in functors of nouns and permits dispensing with the positing of extra phonetically null nominal items in the lexicon.

In line with traceless analyses of long distance dependencies, the account presented here is proposed as a further step towards a more lean theory of grammar, without the reification of missing elements as actual empty categories.

References

Allegranza, Valerio. 1998a. Determination and Quantification. In Frank Van Eynde and Paul Schmidt (eds.), *Linguistic Specifications for Typed Feature Structure Formalisms*, pages 281–314, Luxembourg: Office for Official Publications of the European Communities.

Allegranza, Valerio. 1998b. Determiners as functors: NP structure in Italian. In Sergio Balari and Luca Dini (eds.), *Romance in Head-driven Phrase Structure Grammar*, volume 75 of *CSLI Lecture Notes*, pages 55–108, Stanford: CSLI Publications.

Beavers, John. 2003. More Heads and Less Categories: A New Look at Noun Phrase Structure. In Stefan Müller (ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pages 47–67, Stanford: CSLI Publications.

Copestake, Ann, Flickinger, Dan, Sag, Ivan A. and Pollard, Carl. 2005. Minimal Recursion Semantics: An Introduction. *Journal of Research on Language and Computation* 3(2–3), 281–332.

Ginzburg, Jonathan and Sag, Ivan A. 2001. *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. Stanford, California: CSLI Publications.

Hankamer, Jorge and Sag, Ivan. 1976. Deep and Surface Anaphora. *Linguistic Inquiry* 7(3), 391–426.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Lobeck, Anne. 1995. *Ellipsis – Functional Heads, Licensing, and Identification*. New York, Oxford: Oxford University Press.

Masullo, José Pascual. 1999. Variable vs. Intrinsic Features in Spanish Nominal Ellipsis.

Müller, Stefan. 1996. The Babel-System–An HPSG Prolog Implementation. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, pages 263–277, London.

Müller, Stefan and Kasper, Walter. 2000. HPSG Analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 238–253, Berlin Heidelberg New York: Springer-Verlag, Artificial Intelligence edition.

Nerbonne, John and Mullen, Tony. 2000. Null-headed nominals in German and English. In Frank van Eynde, Ineke Schuurman and Ness Schelkens (eds.), *Proc. of Computational Linguistics in the Netherlands 1998*, pages 143–64.

Netter, Klaus. 1996. *Functional Categories in a an HPSG for German*, volume 3 of *Saarbrücken Dissertations in Computational Linguistics and Language Technology*.

Pollard, Carl and Sag, Ivan. 1994. *Head-driven phrase structure grammar*. Chicago University Press and CSLI Publications.

Pullum, Geoffrey K. 1975. *People* Deletion in English. In *Working Papers in Linguistics*, volume 14, pages 95–101, Ohio State University.

Sag, Ivan A. 2000. Rules and Exceptions in the English Auxiliary System. Manuscript.

Ticio, M. Emma. 2005. NP-Ellipsis in Spanish. In David Eddington (ed.), *Selected Proceedings of the 7th Hispanic Linguistics Symposium.*, Somerville, MA.

Van Eynde, Frank. 2003a. On the Notion 'Determiner'. In Stefan Müller (ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pages 391–396, Stanford: CSLI Publications.

Van Eynde, Frank. 2003b. Prenominals in Dutch. In Jong-Bok Kim and Stephen Wechsler (eds.), *The Proceedings of the 9th International Conference on HPSG*, Stanford University: CSLI Publications.

Winhart, Heike. 1997. Die Nominalphrase in einem HPSG-Fragment des Deutschen. In Erhard Hinrichs, Walt Detmar Meurers, Frank Richter, Manfred Sailer and Heike Winhart (eds.), *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, Chapter 5, pages 319–384, Tübingen: Universität Tübingen.