Noun Ellipsis without Empty Categories

António Branco
Universidade de Lisboa

Francisco Costa
Universidade de Lisboa

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 -]

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

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
old men with hat have
[junge - mit Mütze] getroffen.
[young - with cap] met

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

In some languages, like English, definite articles cannot alone form an NP (English example from [Lobeck, 1995]), while others, like German, allow it:

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

b. Wir haben einen Film gesehen
we have a movie watched
aber [der -] war langweilig.
but [the -] was boring

In languages like Spanish and Portuguese, with pre- and post-head adjectives, pre-head ones (which are intensional) cannot appear in this construction ([Ticio, 2005]), although post-nominal adjectives (intensional or not) can:

(5) a. * Ayer vi a la verdadera
yesterday I saw the true
terrorista y a [la supuesta -]
terrorist and [the alleged -]

b. [um - real] e [outro - imaginado]
[a/one - real] and [another - imagined]

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]),

1

(6) a. die   starke   Konzentration   auf die
       the   strong   concentration   on the
       Wirtschaft   und   [die   weniger
       economy   and   [the   less
       grosse -   auf den   Umweltschutz]
       large -   on the   environment

   b. * Juan   visitó   a   sus   tíos
       Juan   visited   his   uncles.MASC
       y   Pedro   visitó   a [la   - suya].
       and   Pedro   visited [the   - his.FEM]

but not necessarily number:

(7) a. Juan   visitó   a   sus   tíos   y
       Pedro   visitó   a[l - suyo].
       Pedro   visited   [the - his.MASC.SG]

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 *people* deletion/null-N generics (e.g., *the desperately poor*), which tend to be limited to descriptions of people and do not resort to an antecedent to be interpreted ([Pullum, 1975], [Nerbonne and Mullen, 2000]).

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.

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

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].

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 also 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, however, how the use of unary schemata has been enhanced since then,[1] this is clearly an option worth considering, and it is the approach that will be explored in the next Sections.

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, 1998] and [Van Eynde, 2003], among others, who propose the simplification resulting from replacing the specific constructs used to handle specifiers and adjuncts by a more general one for functors.[2] 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

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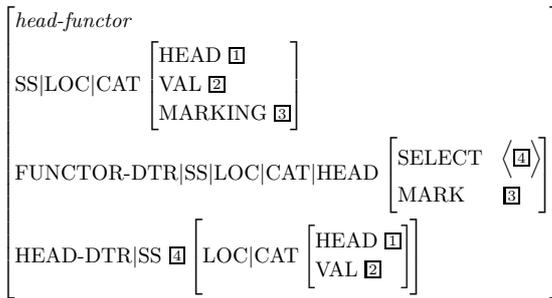
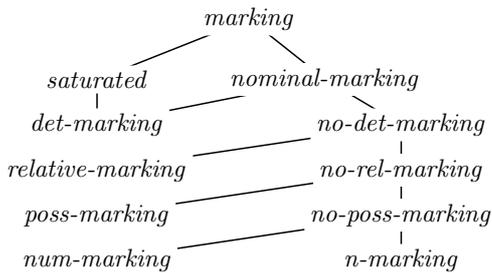Figure 1: Outline of head-functor schema

Figure 2: Type hierarchy of marking

and turns out to require fewer syntactic schemata.

A head-functor schema comes then into play, which, as outlined in Fig. 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.

Under this approach, a major difference between specifiers and adjuncts is that the latter will not change MARKING information, so that they can recur.

A type hierarchy for its possible values can then be used to enforce the subtleties of the NP internal syntactic structure. For the sake of illustration, a type hierarchy for MARKING values that mirrors a simplified NP structure for Portuguese is presented in Fig. 2.[3]

Within this setup, items that select for NPs can then constrain them to be [MARKING *saturated*] (instead of requiring them to have an empty specifier). Possessives select nominals with [MARKING *no-poss-marking*] and impose a [MARK *poss-marking*]. By the same token, nu-

merals select nominals with [MARKING *n-marking*] and have [MARK *num-marking*]; adjective and preposition phrases select for [MARKING *n-marking*] and have [MARK *n-marking*], and so on and so forth. Appropriate nesting is thus enforced with the help of MARKING values: for instance, possessives are more peripheral than numerals.

Word order between functor and head must be enforced separately.[4] Notice that the use of a unary schema can also be assumed to produce bare NPs when appropriate.

Therefore, in this approach, 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 constituents they yield when they are attached to their head (under the MARK feature). This account of NPs in general brings the important advantage that 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.

4 Analysis

Against this background, elliptical NPs without complements can be easily accounted for with the help of a syntactic schema *ellip-head-functor* which is a straightforward unary version of the schema in Fig. 1 for NPs but without the HEAD-DTR. This schema relies on the selectional information made available by the functor about its head, in SELECT, to make the properties of the (missing) noun operative:

- 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 in what concerns the HEAD and VAL feature values.[5] As for the remainder features, note that, on the one hand, the SYNSEM|LOCAL|CONT|RELS of the mother node must be the union of the functor's RELS with

---

[3]To the best of our knowledge, no attempts have been made yet to establish hierarchies for *marking* that can be assumed as universal. The one we presented above does not bear that claim either, though it does match the NP structure widely assumed in the transformational literature.

[4]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.

[5]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.

a singleton set with a nominal object in it (a *noun-ellipsis* relation[6]); on the other hand, like in the example with possessives above, the MARKING values 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.

Example: We present an example parse of the Portuguese NP *os meus dois* (lit. *the my two*, i.e. *my two*). The resulting parse tree is in Fig. 3. The numeral *dois* feeds the *ellip-head-functor* rule and yields a node with [HEAD *noun*] and [MARKING *num-marking*]. The possessive attaches normally, via some head-final version of the head-functor schema in Fig. 1, producing a node with [MARKING *poss-marking*]. The article is merged in the same way, producing a full (saturated) NP. The resulting structure is completely parallel to the one of an NP like *os meus dois carros* (*my two cars*), except for the branch connecting to the missing N node.

The NP *os meus* will also be generated (with *meus* feeding the ellipsis rule), but then the numeral cannot attach due to conflicting values of MARKING (and to other constraints that also block numerals from attaching to the right of nouns in Portuguese).

Structural Parallelism: Maintaining structural parallelism between NPs with expressed nouns and NPs with null nouns has several advantages. For instance, if we assume that relative clauses attach lower than determiners in expressed noun NPs, in the present analysis an NP like *some that arrived* will get the structure [[*some$_D$*] [[*that arrived$_{RC}$*] $_{\bar{N}}$] $_{NP}$]. 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. This analysis has problems attaching a relative clause (or a numeral as in *these three*, etc.) to an unexpressed complement and will thus be forced to have these elements attached to full NPs when no noun is present. In the analysis presented here, determiners, with [MARK *det-marking*], produce full NPs when they undergo the ellipsis rule, but if a relative clause is present, it is this element that feeds *ellip-head-functor* and the determiner attaches higher.

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 Principle B binding ([Hankamer and Sag, 1976], [Lobeck, 1995] among others; the following examples are theirs). In fact, the antecedent can be given pragmatically, as in (9.a), or be in a different sentence (9.b).

(9) 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 (3.a) illustrates this point). We will continue using the expression *one anaphora* in this text though.

Complements of null nouns: The COMPS feature of the node produced by the *ellip-head-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:[7]

---

[6]This is a temporary solution because we are not recovering the semantics of the missing noun. See the discussion on antecedent resolution below.

[7]This is why in *ellip-head-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 ad-nominals 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 (10).
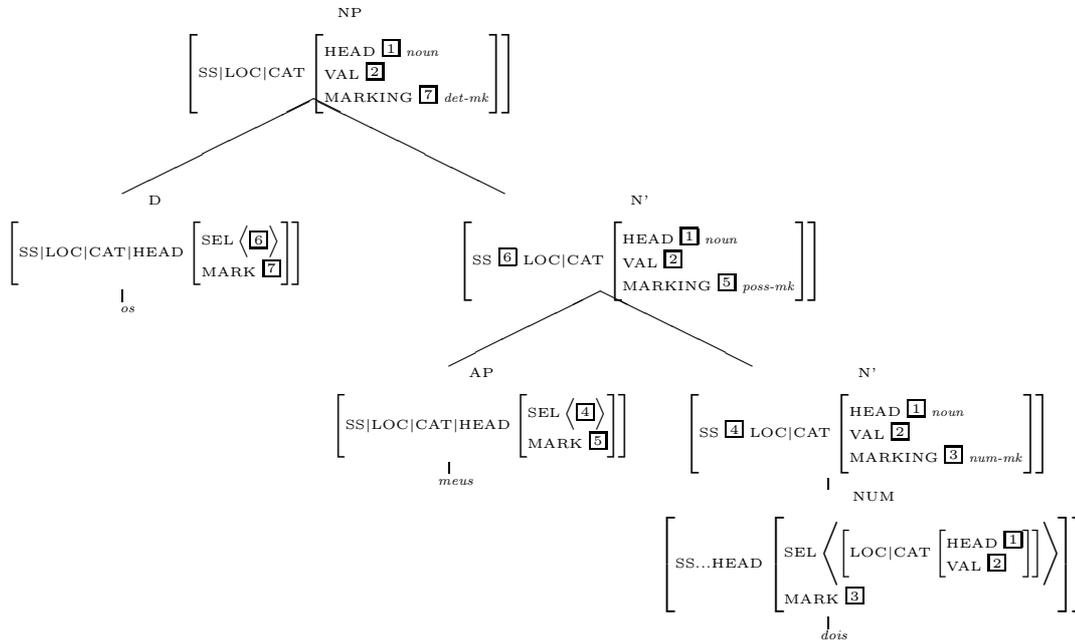
NP

$$\left[\text{SS|LOC|CAT} \begin{bmatrix} \text{HEAD} \ \boxed{1} \ noun \\ \text{VAL} \ \boxed{2} \\ \text{MARKING} \ \boxed{7} \ det\text{-}mk \end{bmatrix} \right]$$

D

$$\left[\text{SS|LOC|CAT|HEAD} \begin{bmatrix} \text{SEL} \ \langle \boxed{6} \rangle \\ \text{MARK} \ \boxed{7} \end{bmatrix} \right]$$

os

N'

$$\left[\text{SS} \ \boxed{6} \ \text{LOC|CAT} \begin{bmatrix} \text{HEAD} \ \boxed{1} \ noun \\ \text{VAL} \ \boxed{2} \\ \text{MARKING} \ \boxed{5} \ poss\text{-}mk \end{bmatrix} \right]$$

AP

$$\left[\text{SS|LOC|CAT|HEAD} \begin{bmatrix} \text{SEL} \ \langle \boxed{4} \rangle \\ \text{MARK} \ \boxed{5} \end{bmatrix} \right]$$

meus

N'

$$\left[\text{SS} \ \boxed{4} \ \text{LOC|CAT} \begin{bmatrix} \text{HEAD} \ \boxed{1} \ noun \\ \text{VAL} \ \boxed{2} \\ \text{MARKING} \ \boxed{3} \ num\text{-}mk \end{bmatrix} \right]$$

NUM

$$\left[\text{SS...HEAD} \begin{bmatrix} \text{SEL} \ \left\langle \left[\text{LOC|CAT} \begin{bmatrix} \text{HEAD} \ \boxed{1} \\ \text{VAL} \ \boxed{2} \end{bmatrix} \right] \right\rangle \\ \text{MARK} \ \boxed{3} \end{bmatrix} \right]$$

dois

Figure 3: Example NP with a null noun. The NP is [*os meus dois* -]; lit. [*the my two* -], i.e. [*my two* -]. The structure is parallel to the one assumed for [$_{NP}$ *os* [$_{\bar{N}}$ *meus* [$_{\bar{N}}$ *dois* [$_{\bar{N}}$ *carros*]]]] (*my two cars*).

(10) o   filho   mais   velho   do   Pedro
the   son   most   old   of the   Pedro
e   o   [$_{\bar{N}}$ [$_N$ -   mais novo]   [da Maria]]
and   the   -   most young   of the Maria
Pedro's eldest son and Maria'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, *ellip-head-comp* is required. For example, we want to assign to the NP *o da Maria* the structure presented in (11), but the most embedded PP is not a functor.

(11) o   filho   do   Pedro
the   son   of the   Pedro
e   [$_{NP}$ o   [$_{\bar{N}}$ -   [$_{PP}$da   Maria]]]
and   the   -   of the   Maria
Pedro's son and Maria'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 normally. This rule should of course ensure that the complement PP is compatible with the selectional properties inherited from the antecedent of the elliptical NP.

Some of the key properties of this schema are common to the *ellip-head-functor* schema above. The resulting constituent:

- has a HEAD of type *noun*;

- 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;

- its RELS value is the union of the complement daughter RELS value with a singleton set in which there is a *noun-ellipsis* relation.

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 *ellip-head-functor* may suffice to account for all data concerning noun ellipsis.[8]

We now turn to the specific co-occurrence restrictions inside elliptical NPs.

---

[8]Additionally, the COMPS of the mother node in *ellip-head-functor* must be constrained to be empty in these languages.

5

Edge Features: The *Left Periphery (Empty)* (LP(E)) attribute that is used in some analyses ([Nerbonne and Mullen, 2000], [Netter, 1996])[9] to control many co-occurrence restrictions in elliptical NPs is compatible with the present proposal and can be kept in use for the same purpose.

In order to make use of LP in this approach, one needs to split each ellipsis rule into two rules: one that corresponds to the case where the missing noun precedes the functor or complement, the other to the case where it follows it. Note that not all combinations will be necessary for every language, as in languages with fixed head-complement word order only one *ellip-head-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* — see the last paragrah in this block). 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.

An additional constraint must be added to *ellip-head-functor* rules — their functor daughter selects an element with [LP *empty*] (since it has no sister node):

$$\left[\text{FUNCTOR-DTR|SS|LOC|CAT|HEAD|SEL} \left\langle \left[\text{LP } empty \right] \right\rangle \right]$$

Of course, in the remaining schemata, LP must percolate from the leftmost daughter, and lexical items (other than the anaphoric *one*, which is [LP *one*]) are [LP *full*].

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

***One* anaphora vs. missing nouns:** The LP feature that has been 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. [10]

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, their contribution for an empirically correct account of noun ellipsis can be put to work also in this case (the definite articles can select an element with [LP ¬*empty*]).[11]

Pre-head adjectives: In some languages like Spanish or Portuguese, predominantly with post-head adjectives, pre-head adjectives, of the *alleged* type, 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*].[11]

5 Related Issues and Open Questions

An alternative analysis to control the interaction of elliptical NPs with definite articles and obligatorily prenominal adjectives is presented in [Costa, to appear]. It correlates non-eligibility to feed the noun ellipsis rules with the lack of complements of these adjectives (and also the articles, although it is not clear what their complements might be) and the requirement of occurring as pre-nominals. It is thus less stipulative than the LP constraints, but lack of space precludes a presentation of this analysis here.

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

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

Note that the difference between the two constructions also involves lexical idiosyncrasies. For instance, NPs with numerals feeding the ellipsis/null-N generics rules (e.g. *the two*) do not allow null-N generic readings, but surprisingly in English they do not co-occur with *one*. Future research may use such considerations to shed light on the distribution of *one* (only used in

---

the languages that have them when there would otherwise be potential ambiguity between the two constructions?), which is accounted for by the LP analysis in a completely stipulative fashion.

Another issue is the status of personal pronouns. Under the present analysis the question arises of whether personal pronouns are fully saturated NPs or rather determiners occurring in noun-less NPs. Phrases like the English *you two* or *we students* suggest relevant plausibility for the latter, but their correlates in other languages do not confirm it straightforwardly: e.g. Portuguese *vocês os dois* (lit. *you the two*) and *nós os estudantes* (lit. *we the students*).

6 Concluding remarks

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.

7 Acknowledgments

References

Allegranza, V. 1998. Determination and quantification. In Van Eynde, F. and Schmidt, P., editors, *Linguistic Specifications for Typed Feature Structure Formalisms*, pages 281–314. Office for Official Publications of the European Communities, Luxembourg.

Beavers, J. 2003. More heads and less categories: A new look at noun phrase structure. In Müller, S., editor, *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pages 47–67, Stanford. CSLI Publications.

Costa, F. to appear. Noun-less noun phrases in a deep parser of Portuguese. Master's thesis, Universidade de Lisboa.

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

Hankamer, J. and Sag, I. 1976. Deep and surface anaphora. *Linguistic Inquiry*, 7(3):391–426.

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

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

Masullo, J. P. 1999. Variable vs. intrinsic features in Spanish nominal ellipsis.

Nerbonne, J. and Mullen, T. 2000. Null-headed nominals in German and English. In van Eynde, F., Schuurman, I., and Schelkens, N., editors, *Proc. of Computational Linguistics in the Netherlands 1998*, pages 143–64.

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

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

Sag, I. A. 2000. Rules and exceptions in the English auxiliary system. Manuscript.

Ticio, M. E. 2005. NP-Ellipsis in Spanish. In Eddington, D., editor, *Selected Proceedings of the 7th Hispanic Linguistics Symposium.*, Somerville, MA.

Van Eynde, F. 2003. Prenominals in Dutch. In Kim, J.-B. and Wechsler, S., editors, *The Proceedings of the 9th International Conference on HPSG*, Stanford University. CSLI Publications.

Winhart, H. 1997. Die Nominalphrase in einem HPSG-Fragment des Deutschen. In Hinrichs, E., Meurers, W. D., Richter, F., Sailer, M., and Winhart, H., editors, *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, chapter 5, pages 319–384. Universität Tübingen, Tübingen.