# The syntax of French N′ phrases

Anne Abeillé
Universite Paris 7

Olivier Bonami
Universite Paris 7

Danièle Godard
Universite Paris 7

Jesse Tseng
Universite Paris 7

Proceedings of the 11th International Conference on
Head-Driven Phrase Structure Grammar

Center for Computational Linguistics, Katholieke Universiteit Leuven

Stefan Müller (Editor)

2004

Stanford, CA: CSLI Publications

pages 6–26

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2004. The syntax of French N′ phrases. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 6–26. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2004.1.

**Abstract**

Of all French functional elements, the form *de* has without question the widest variety of uses, and presents the greatest challenge for linguistic description and analysis. Historically a preposition, it still has a number of prepositional uses in modern French, but in many contexts it calls for an altogether different treatment. We begin by outlining a general distinction between "oblique" and "non-oblique" uses of *de*. We then develop a detailed account of constructions where *de* combines with an N′. We provide a unitary analysis of *de* in three constructions (quantifier extraction, "quantification at a distance", and negative contexts) which have been not been considered to be related in previous accounts.

The aim of this article is to present a novel and unified analysis of structures of the form [*de* N′] in French. After giving an overview of the uses of the element *de* and establishing a partition of these uses corresponding to two distinct syntactic analyses for *de* (section 1), we provide a detailed description of the *de*-N′ structures examined in the rest of the paper (section 2). Section 3 presents arguments for treating certain *de*-N′ phrases as extraction sites, and section 4 provides the full HPSG analysis, with example structures.

# 1 Background: The dual syntactic nature of *de*

This section briefly motivates a two-way classification of the uses of *de* based on a number of syntactic criteria, and presents an HPSG account of the data. For a more complete presentation (including similar results for the element *à*), see Abeillé et al. (2003).[1]

## 1.1 Oblique vs. non-oblique uses

First we can identify "oblique" uses of *de*, characterized by the following properties: nothing can be extracted from the phrase that *de* combines with (1), *de* can combine with a coordination of phrases (2), and the *de*-phrase cannot appear in subject position (3).

(1) Je me souviens [de la fin de ce film].
'I remember the end of that film.'

⇝ *un film dont je me souviens [de la fin ___ ]
'a film of which I remember the ending'

(2) J'ai besoin [ de [cette farine et cette levure]] .
'I need this flour and this yeast.'

†We would like to thank the participants of the group "Grammaire typologique des formes faibles" at Paris 7 University, in particular Denis Creissels and Anne Zribi-Hertz, and the audience of the HPSG Conference in Leuven, in particular Henriëtte de Swart and Frank Van Eynde, for their valuable comments on earlier presentations of this work.

[1]Kupferman (2004) presents a related approach to partitioning the uses of *de*.

(3) *[De mort] est la seule façon efficace de menacer ces gens.
'With death is the only effective way to threaten these people.'

Using these tests, we find that oblique *de* appears in combination with NP and N′ (as in the preceding examples), and also with PP, AP, and AdvP (but never with VP):

(4) a. Il surgit [de derrière le rideau].
'He jumped out from behind the curtain.'
b. quelque chose [de plus traditionnel]
'something more traditional'
c. deux jours [de plus]
'two days more'

In "non-oblique" uses, *de* behaves very differently, allowing extraction out of its sister phrase (5), not taking wide scope over a coordination of phrases (6), and sometimes occurring in subject position (7).

(5) a. Je n'ai pas lu [de livres de cet auteur].
'I haven't read any books by this author.'
⇝ un auteur dont je n'ai pas lu [de livres ___ ]
'an author of whom I haven't read any books'
b. Je ne me souviens pas [d'avoir lu ce livre].
'I don't remember having read that book.'
⇝ un livre que je ne me souviens pas [d'avoir lu ___ ]
'a book I don't remember having read'

(6) *On nous a apporté plein ⌊ de [pain et vin]⌋ .
'They brought us loads of bread and wine.'

(7) [De sortir un peu] te ferait du bien.
'Getting out a bit would do you some good.'

In addition to these examples containing N′ and infinitival VP, non-oblique *de* also combines with the definite article to give rise to the so-called "partitive article" (8a). This construction is in fact available with other demonstrative and possessive specifiers as well (8b) (Kupferman, 2004).

(8) a. un courrier contenant [de la poudre blanche suspecte]
'a letter containing suspicious white powder'
b. acheter [de ce/son whisky]
'buy some of that/his kind of whisky'

## 1.2 Analysis

The properties of oblique *de*-phrases can be accounted for in a straightforward manner if we analyze *de* as an ordinary preposition, satisfying the lexical description in (9). In French, PPs are extraction islands, hence the empty SLASH list for

all *prep-words* (10b).[2] And unlike their English counterparts, French PPs cannot be used as subjects (3,11).

(9) prepositional *de*: *prep-word* &

$$\begin{bmatrix} \text{MARKING} & \textit{de} \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & \neg\, \textit{verb} \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

(10) a. *une loi dont j'ai voté pour l'auteur
'a law of which I voted for the author (whose author I voted for)'

b. $\textit{prep-word} \rightarrow \begin{bmatrix} \text{HEAD} & \textit{prep} \\ \text{MARKING} & \textit{marked} \\ \text{SLASH} & \{\,\} \end{bmatrix}$

(11) *Sous la table est une bonne cachette.
'Under the table is a good hiding place.'

Non-oblique *de*-phrases, on the other hand, do not behave like PPs, but more like NPs or VPs. We propose that non-oblique *de* is a "weak head"—that is, a syntactic head that shares its HEAD value with its complement (Tseng, 2002). One lexical entry for the weak head *de*, used with nominal and verbal complements, is shown in Figure 1.

$$\begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{MARKING} & \textit{de} \\ \text{SUBJ} & \boxed{2} \\ \text{SPR} & \boxed{3} \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{MARKING} & \textit{unmarked} \\ \text{SUBJ} & \boxed{2} \\ \text{SPR} & \boxed{3} \\ \text{COMPS} & \langle\,\rangle \\ \text{CONT} & \boxed{4} \end{bmatrix} \right\rangle \\ \text{CONT} & \boxed{4} \end{bmatrix}$$

Figure 1: Weak head *de* #1

As an example, if the weak head *de* selects a verbal complement as in (7), the resulting [*de* VP] combination has the HEAD value [*verb*, VFORM *inf*] and

[2] This observation holds for standard varieties of European French. In other varieties, the constraint can be relaxed in certain contexts.

we expect it to have the properties of an infinitival VP with respect to distribution, extraction, and so on. In this case, *de* will also inherit the non-empty SUBJ list of its VP complement, which will then be visible on the dominating phrase (crucial for the analysis of raising and control). Note that in our analysis, all [*de* VP] structures involve the weak head of Figure 1, and never the preposition in (9).[3]

On the other hand, if the weak head *de* combines with a nominal complement as in *plein de pain* (6), the resulting [*de* NP] combination is correctly predicted to have the grammatical properties of an NP. Note that for "partitive" NPs as in (8), a distinct weak head entry for *de* is required, one that selects an NP complement introduced by a definite, demonstrative, or possessive specifier, and contributes the appropriate partitive/indefinite semantics.

Oblique and non-oblique *de* do share one crucial property: the MARKING value *de*, which then propagates to the phrases they project. This explains why all pronominalizable *de*-phrases alternate with the clitic *en*, despite their otherwise highly divergent grammatical properties (12). The principles governing *en*-cliticization refer only to the feature [MARKING *de*].

(12) a. Je me souviens [de ce film]. *(oblique)*
'I remember that film.'
$\rightsquigarrow$ Je m'en souviens. 'I remember it.'

b. Je n'ai pas lu [de livres]. *(non-oblique)*
'I haven't read any books.'
$\rightsquigarrow$ Je n'en ai pas lu. 'I haven't read any.'

The MARKING feature is also used to prevent iteration of the weak head *de*. As indicated in Figure 1, its complement must be *unmarked*, and so cannot already be a projection of *de*. Another consequence of the *unmarked* constraint is that the weak head cannot select a prepositional complement, because all prepositions introduce a *marked* specification, as shown in (10b). Prepositional *de* (9), on the other hand, can take a *marked* PP complement, as in (4a), or even a [MARKING *de*] complement headed by the weak head *de*.[4]

(13) J'ai besoin de [beaucoup de farine].
'I need a lot of flour.'

[3] This is in contrast, for example, to Huot 1981, who proposes either a PP or a VP analysis depending on the higher verb.

[4] See section 4.1 for details of the analysis of the bracketed *de*-phrase. Note that prepositional *de* cannot be immediately followed by another *de*:

(i) *J'ai besoin de [de la farine].
'I need (some) flour.'

(ii) J'ai besoin de farine.

A full account of these "cacophony" effects (Gross, 1967) prohibiting adjacent occurrences of *de* must incorporate constraints referring to linear word order (e.g., the notion of "left edge"). The grammatical alternative to (i) is the so-called "haplology" construction (ii) with a single occurrence of *de* (and no definite article). A special prepositional entry for *de*, selecting an $N'$ complement (with the appropriate semantics), is required for such examples.

*De* also displays some morphophonological idiosyncrasies: it always undergoes vowel elision (*de* $\rightsquigarrow$ *d'*) conditioned by the following context, and contraction with the specifier forms *le* and *les* (giving rise to *du* and *des* respectively). The oblique vs. non-oblique status of *de* has no influence on its morphophonological behavior. Finally, we note that the partition into prepositional (oblique) and weak head (non-oblique) uses proposed here does not correlate with any semantic criteria. In particular, there are semantically empty prepositional uses of *de* and semantically potent weak head uses.

## 2 Nominal *de*-phrases

In the remainder of this paper, we focus on the various types of (non-oblique) nominal phrases of the form *de*-N′, which have restricted distribution and must always be licensed by other material. In (14), for instance, the *de*-phrase is not licensed, and the sentence is ungrammatical.

(14) *J'ai lu [de livres].
'I read DE books.'

There are several ways in which example (14) can be extended to produce a grammatical sentence.[5]

### 2.1 Local quantification

In the simplest case, *de*-N′ can be licensed locally by a degree expression from a class including adverbs (*beaucoup*, *infiniment*, *combien*), nouns (*nombre*, *quantité*), or the invariable form *plein*. The resulting phrases of the form [Deg *de* N′] (e.g., *beaucoup de livres* 'a lot of books') have the distributional properties of ordinary NPs: they can appear as subject or complement of a verb, or as complement of a preposition.

(15) a. Nous avons perdu [beaucoup de livres].
'We lost a lot of books.'
b. [Beaucoup de livres] ont été abîmés.
'A lot of books were damaged.'

[5] Cases not considered here include:

(i) *de* as allomorphic variant of partitive *des* before pre-nominal modifier
J'ai lu [de très vieux livres].
'I read some very old books.'

(ii) dislocated *de*-N′ (Milner, 1978)
J'en veux trois, de robes.
'I want three dresses.'

(iii) rare occurrence in some negative polarity contexts (Gaatone, 1971, 1992; Muller, 1997)
A-t-on jamais publié de livre aussi mauvais ?
'Has such a bad book ever been published?'

c. Il est parti avec [beaucoup de livres].
'He left with a lot of books.'

Semantically, we have a mass/plural nominal expression with the degree element functioning as an intersective quantifier. For example, in *lire beaucoup de livres*, the quantity of books read is measured against some contextually determined scale and found to be 'a lot'.

Note that a complete analysis of degree adverbs should be able to relate this use to occurrences of the same adverbs as verb modifiers—see (17a) below, for example. Abeillé and Godard (2003) propose treating degree adverbs uniformly as modifiers, even in the [Deg *de* N′] construction. While such an analysis is intuitively appealing, it faces problems at the syntax-semantics interface: [Deg *de* N′] has scopal properties typical of a quantified NP, as illustrated by the interaction with negation in (16). Thus the degree expression in [Deg *de* N′] does not behave semantically like a noun modifier, but more like a specifier.

(16) a. Paul n'a pas emprunté beaucoup de livres.
Most salient reading: 'It is not the case that Paul borrowed a large number of books.' (NEG $>$ *beaucoup*)

b. Beaucoup de livres n'ont pas été empruntés par Paul.
Most salient reading: 'There is a large number of books $x$ such that it is not the case that Paul borrowed $x$.' (*beaucoup* $>$ NEG)

A more adequate way to relate the ad-verbal and ad-nominal uses is to assume that the relation between an entity and a scale associated with a degree expression is used to form an intersective modifier in the ad-verbal use, and to specify the size of the group which is also quantified over in the [Deg *de* N′] construction. In the semantic forms below, $\mathcal{S}$ represents a contextually-supplied standard degree scale.

(17) a. Paul a beaucoup dormi.
'Paul slept a lot.'
$\exists e[\mathsf{sleep}(e, \mathsf{p}) \wedge \mathsf{a\text{-}lot}(e, \mathcal{S})]$

b. Beaucoup d'enfants dorment.
'Many children are sleeping.'
$\exists e \exists X[\mathsf{children}(X) \wedge \mathsf{sleep}(e, X) \wedge \mathsf{a\text{-}lot}(X, \mathcal{S})]$

For the sake of brevity, we will treat the degree elements in [Deg *de* N′] phrases as atomic binary quantifiers.

## 2.2 Non-adjacent quantifier

There are two variants of *de*-N′ licensed by degree quantification in which the degree element does not form a constituent with the *de*-N′ phrase.

First, the licensing adverbial can be the filler in a filler-gap construction: either the interrogative *wh*-word *combien* 'how much/how many' (18a), or correlative

*plus*, *moins* 'more, less'.[6]

(18) Quantifier extraction

a. Combien as-tu lu [___ de livres en latin] ?
how many have you read DE books in Latin
'How many books have you read in Latin?'

b. Plus Paul veut lire [___ de livres], plus il va à la
more Paul wants read DE books more he goes to the
bibliothèque.
library
'The more Paul wants to read books, the more he goes to the library.'

The split *combien* ... *de*-N′ construction constrasts with interrogatives containing the NP [*combien de* N′], either in situ or in filler position:

(19) a. Tu as lu [combien de livres en latin] ?
'You've read how many books in Latin?'

b. [Combien de livres en latin] as-tu lus ?
'How many books in Latin have you read?'

Second, a subset of the degree expressions found in [Deg *de* N′] structures (*beaucoup*, *trop*, *assez*, ...) can "float" immediately to the left of an infinitive or past participle, giving rise to "quantification at a distance" (henceforth QAD).

(20) Il va [beaucoup lire [de livres]] / Il a [beaucoup lu [de livres]].
he goes MANY read DE books / he has MANY read DE books
'He's going to read many books / He has read many books.'

In fact the degree adverb cannot be arbitrarily distant from the *de*-N′ phrase it licenses; QAD is VP-bounded (21).

(21) Paul m'a forcé $[_{\mathrm{VP}}$ à boire [beaucoup de pastis]$]$ .
'Paul forced me to drink a lot of pastis'
$\rightsquigarrow$ *Paul m'a beaucoup forcé $[_{\mathrm{VP}}$ à boire [de pastis]$]$ .

In contrast to the ambiguity observed for [Deg *de* N′] phrases (22a)[7] the adverb in QAD systematically takes narrow scope (22b).

(22) a. Paul recevra chaque étudiant qui a lu [beaucoup de livres]. (local Q, two readings)
(i) 'Paul will meet every student who read a large number of books' *or*
(ii) 'There is a large number of books $x$ such that Paul will meet every student who read $x$'

[6] See Borsley, 2004 for arguments in favor of treating correlative constructions as filler-gap structures.

[7] Recall also the the examples in (16). It should be noted that the wide scope reading (ii) in (22a) is not accepted by all speakers.

b. Paul recevra chaque étudiant qui a beaucoup lu [de livres]. (QAD, reading (i) only)

Moreover, QAD is not semantically compatible with all predicates (23). According to Obenauer (1994) and Doetjes (1997), this indicates that the adverb quantifies over the predicate, and only indirectly over the nominal argument.

(23) a. Jean a vu / apprécié [beaucoup de films]. (local Q)
'Jean has seen/appreciated many films'

b. Jean a beaucoup vu / ??apprécié [de films]. (QAD)

The data are quite tricky, however, and it is unclear how this general proposal can be implemented in a fully explicit semantics. For the purposes of this paper we treat QAD as quantification over individuals.

## 2.3 Negative contexts

Finally, *de*-N′ phrases can be licensed by negation.[8]

(24) Paul n'a pas lu [de livre].
'Paul did not read any book.'

In such cases *de*-N′ is interpreted as an existential quantifier in the immediate scope of the negation. This is in contrast to the ambiguity of example (25) involving the indefinite article *un*; sentence (24) has only the interpretation (i).

(25) Paul n'a pas lu un livre.
(i) 'Paul read no book.'
(ii) 'There is a book that Paul didn't read.'

There are a number of non-verbal negative licensers, including the preposition *sans* 'without' + VP[*inf*] (26), and left-adjoined negative adverbs (27).[9]

(26) Il est parti sans donner [d'explications].
'He left without giving any explanations.'

(27) a. Pas [de problème] !
'No problem!'

b. Paul préfère une mauvaise solution à pas [de solution] du tout.
'Paul prefers a bad solution to no solution at all.'

[8]The noun can be singular or plural, with no change in meaning.

(i) Je n'ai pas lu de journal / de journaux.
'I did not read any newspaper / newspapers.'

[9]The preposition *sans* cannot license a *de*-N′ phrase as its own complement; in this case, it takes a bare (*unmarked*) N′ complement:

(i) Il est parti sans explications/*[d'explications].
'He left without explanations.'

# 3 Autonomous *de*-N′ phrases as extraction sites

We propose a unified treatment of "autonomous" *de*-N′ phrases—that is, those that do not form a NP constituent with their licenser (an extracted quantifier, a QAD adverb, or negation).[10] Our approach is motivated by the following observations.

## 3.1 Distribution

The *de*-N′ can be a direct object (see examples (18), (20), and (24)), but never a pre-verbal subject or complement of a preposition (28–29). With quantifier extraction and negation, *de*-N′ phrases can also be licensed in post-verbal (inverted) subject position (30).

(28) a. *Combien dis-tu que [de clients] sont venus ?
'How many clients do you say came?'

b. *[De clients] sont beaucoup venus ce matin.
'Many clients came this morning.'

c. *[D'enfants] ne vont pas là-bas.
'No children go there.'

(29) a. *Combien as-tu voté $[_{\mathrm{PP}}$ contre [de projets]] ?
'How many projects did you vote against?'

b. *Paul a beaucoup voté $[_{\mathrm{PP}}$ contre [de projets]] .
'Paul has voted against many projects.'

c. *Paul n'a jamais voté $[_{\mathrm{PP}}$ contre [de projet]] .
'Paul never voted against any project.'

(30) a. Combien dis-tu que sont venus [de clients] ?
'How many clients do you say came?'

b. *un jour où sont beaucoup venus [de clients]
'a day when many clients came'

c. un endroit où ne vont pas [d'enfants]
'a place where no children go'

## 3.2 Unbounded dependency

Extracted quantifiers and negative licensers can be arbitrarily distant from the *de*-N′ phrase:

(31) a. Combien Paul voulait-il [que Marie lise [de livres]] ?
'How many books did Paul want Marie to read?'

b. Paul ne voulait pas [que Marie lise [de livres]].
'Paul did not want Marie to read books.'

QAD, on the other hand, is bounded, as we saw in (21).

[10] See Kayne (1981) for an early proposal along similar lines.

### 3.3 Island constraints

Quantifier extraction and, more surprisingly, licensing by negation obey the PP island constraint (29), the complex NP constraint (32), and the subject constraint (33). We show no results for QAD, which obeys these constraints vacuously, since all of the relevant contexts are already excluded by VP-boundedness.

(32) a. *Combien connais-tu un scientifique [qui a [d'idées sur ce sujet]] ?
'You know a scientist who has how many ideas on this topic?'

b. *Je ne connais pas un scientifique [qui ait [d'idées sur ce sujet]].
'I don't know a scientist who has any ideas on this topic.'

(33) a. *Combien dis-tu que [lire [de livres]] t'a plu ?
'You say that reading how many books pleased you?'

b. *[Que Paul ait lu [de livre]] ne m'a pas surpris.
'That Paul read a book did not surprise me.'

### 3.4 Coordinate structure constraint

Here, the results are less clear, with the CSC strictly enforced in quantifier extraction structures (34), but less so in cases of QAD (35) and negation (36).

(34) a. Combien as-tu [lu de livres] et [feuilleté de magazines] ?
'How many books have you read and how many magazines have you leafed through?'

b. *Combien as-tu [lu de livres] et [feuilleté un magazine] ?
'How many books have you read and leafed through a magazine?'

(35) a. Paul a trop mangé [de pizza] et [de glace].
'Paul has eaten too much pizza and ice cream.'

b. *Paul a trop mangé [de pizza] et [trois glaces].
'Paul has eaten too much pizza and two ice creams.'

c. ?Paul a trop mangé de pizzas et ce genre de glace.
'Paul has eaten pizza, and this type of ice cream, on too many occasions'.

(36) a. Paul n'a pas mangé de gâteau ou de cerises.
'Paul ate neither cake nor cherries.'

b. *Paul n'a pas mangé de gâteau ou la pomme.
'Paul ate neither a piece of the cake nor the apple.'

c. Paul ne veut pas écouter [de disque de Johnny] ou regarder [de film avec lui].
'Paul wants neither to listen to one of J's albums nor to watch a movie with him in it.'

d. ?Paul ne veut pas écouter [de disque de Johnny] ou aller au cinéma ce soir.

'Paul wants neither to listen to one of J's albums nor to go the movies tonight.'

In (35) we see that in QAD, the *de*-N′ phrase can be coordinated with an ordinary NP if the "floating" degree adverb is interpreted as an iterative modifier on the second conjunct. For negation (36), direct coordination of a *de*-N′ with an ordinary NP is impossible, but intervening VP projections can improve grammaticality.

### 3.5 Proposed approach

The data presented above lead us to make the following proposals:

- In quantifier extraction examples (e.g., involving *combien*), the *de*-N′ introduces a SLASH dependency, terminated by the filler-head structure (as standardly assumed).
- In negative contexts also, *de*-N′ introduces a SLASH dependency, that terminates at the node where the negation is retrieved.
- QAD is a more restricted phenomenon (not long-distance), but we can adopt the same basic syntax for the *de*-N′ phrase (including the introduction of a SLASH element) in order to capture the constraints on its distribution.

The CSC facts are not necessarily problematic: the CSC has been argued to be not a general property of extraction, but a property of some filler-gap constructions, conditioned moreover by the discourse relation that links the conjuncts (Kehler, 2002, pp. 101–142). If so, it is not very surprising that the CSC does not apply to *de*-N′ in negative contexts, which do involve SLASH dependencies, but are not filler-gap constructions.

## 4 An HPSG analysis of *de*-N′

In this section, where we present our formal analysis of *de*-N′ licensing, we assume the theory of extraction, quantifier store, and interrogative constructions of Ginzburg and Sag (2001), and the approach to negation at the syntax-semantics interface of de Swart and Sag (2002).

### 4.1 The basic case: [Deg *de* N′]

Phrases in which *de*-N′ is licensed by an immediately adjacent degree element, as in *beaucoup de livres*, are treated as ordinary head-specifier phrases, where the specifier is a quantificational degree expression (Milner, 1978). The relevant lexical entry for *beaucoup* is given in Figure 2. We propose a typical specifier entry, except that *beaucoup* selects (via SPEC) a category bearing the feature [MARKING *de*]. Consequently, *beaucoup* can only combine with an N′ headed by the weak head *de*. This excludes **beaucoup livres*, even though *beaucoup* can be on the SPR

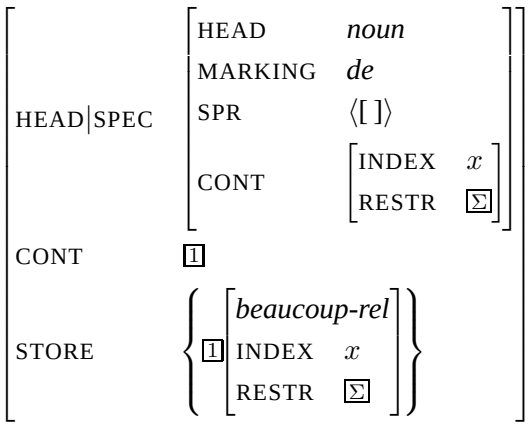


Figure 2: Lexical entry for *beaucoup* (specifier)

list of the common noun *livres*. The lexical entry of the weak head *de* in Figure 1 requires specifier-raising—e.g., the SPR requirement of *livres* becomes the SPR requirement of *de livres*. (Similarly, this entry requires subject-raising in the case of a VP[*inf*] complement.) The full analysis of the phrase *beaucoup de livres* can be seen in Figure 6 at the end of the paper.

### 4.2 *de*-N′ and SLASH

As explained in section 3, we propose treating the licensing of autonomous *de*-N′ phrases in terms of a SLASH dependency. We start by motivating a second lexical entry for the weak head *de*, distinct from the one in Figure 1.

We assume that the first argument of a common noun (i.e., its specifier) must be canonical (37). This accounts for the fact that specifiers of bare N′s cannot be extracted (38).

(37) *cn-wd* $\rightarrow \left[\text{ARG-ST } \langle \textit{canon-ss} \rangle \oplus \textit{list}(\textit{synsem})\right]$

(38) a. Quel livre as-tu lu ? 'Which book did you read?'
b. *Quel as-tu lu livre ?

Given the SPR-list sharing indicated in the weak head entry in Figure 1, the combination of *de* and a common noun (e.g., *de livres*) always gives rise to a SPR-unsaturated nominal projection. We assume that such phrases are disallowed as the direct argument (subject or direct object) of a verb:[11]

(39) *vb-wd* $\rightarrow \left[\text{ARG-ST } \textit{list}\left([\text{SPR } \langle\,\rangle] \vee [\text{PRED } +]\right)\right]$

[11]This constraint does allow predicative nominal arguments to appear without a specifier, as in constructions like *devenir médecin* 'become a doctor' or *faire confiance* 'trust'.

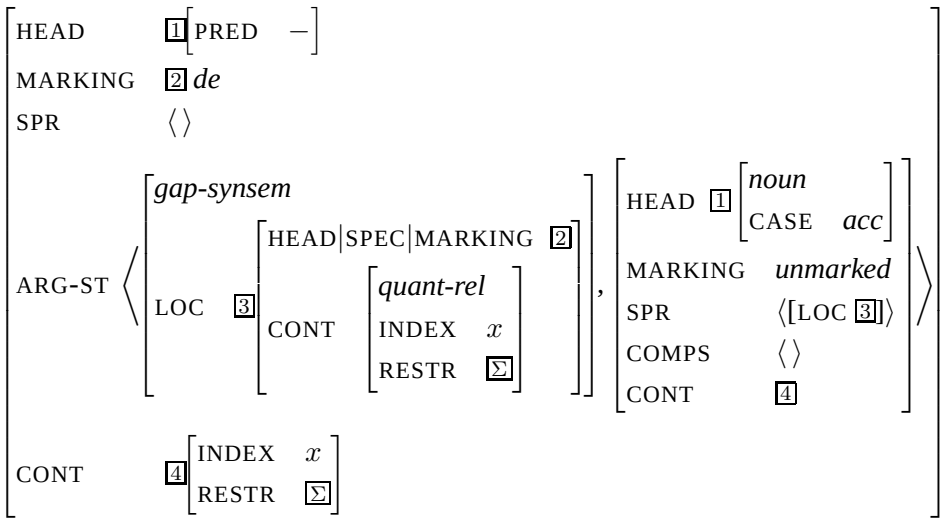


Figure 3: Weak head *de* #2, with slashed specifier

As a consequence of this constraint, a phrase like *de livres*, headed by the weak head in Figure 1, cannot be the direct argument of a verb (unless its SPR-requirement is satisfied by a degree expression like *beaucoup*).[12]

Those *de*-N′ phrases that do occur autonomously must therefore be headed by a different weak head *de*, whose lexical entry is shown in Figure 3. The complement of this weak head (the second item on its ARG-ST list) is an unmarked, SPR-unsaturated nominal. Instead of inheriting the SPR requirement, like the weak head of Figure 1, this variant of *de* has an empty SPR list. This means that a phrase like *de livres*, headed by this weak head, can be used as a direct argument without violating constraint (39). The accusative case specification ensures that these *de*-N′ phrases cannot be preverbal subjects[13] or complements of a preposition (which bear *internal* case, Abeillé and Godard, 1999); recall the examples in (28–29) above. And the feature [PRED −] prevents the occurrence of predicative *de*-N′ phrases in negative contexts.

(40) *Nous n'avons pas été [d'idiots]. 'We were no fools'

The most important aspect of the entry in Figure 3, however, is that the LOCAL value of the N′ complement's unrealized specifier ends up in the SLASH set of *de* (as a result of SLASH Amalgamation applied to the ARG-ST list). The problem of licensing autonomous *de*-N′ phrases thus becomes a matter of formulating the

[12] As it stands, the entry in Figure 1 allows *de* to combine with full NPs, incorrectly producing sentences like **Paul voit de Marie / de trois poissons* ('Paul sees DE Marie / DE three fish'). This can be excluded by further stating that the weak head must inherit a specifier (or a subject); more precisely, either SPR or SUBJ must be non-empty (but not both).

[13] They can appear as inverted subjects (30), which are accusative (Abeillé, 1997; Bonami et al., 1999).

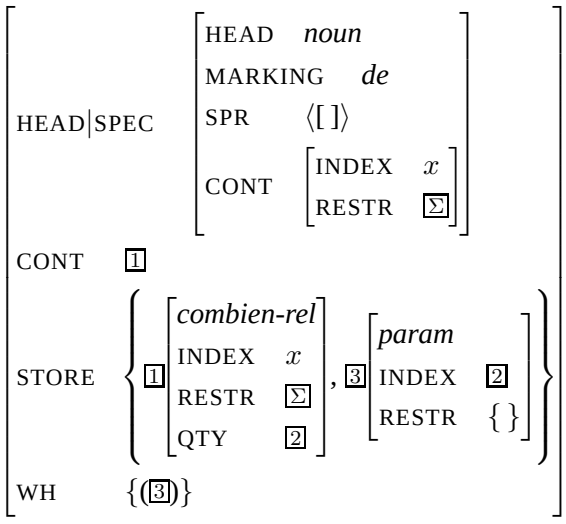


Figure 4: Lexical entry for *combien*

appropriate conditions for discharging the SLASH dependency introduced by the weak head. See Figure 7 at the end of the paper for the structure of the VP *lire de livres*; the SLASH set must somehow be emptied if this VP is to be part of a well-formed sentence.

### 4.3 Quantifier extraction

As an example of our treatment of quantifier extraction—see the data in (18)—we propose the lexical entry in Figure 4 for the specifier *combien*. This entry is syntactically similar to that of *beaucoup* in Figure 2, selecting an N′ with the feature [MARKING *de*] via SPEC; after all, *combien* does appear in [Deg *de* N′] examples like (19), involving the first weak head *de* of Figure 1. In its semantic content, *combien* includes an extra argument for a quantity parameter, which is put on STORE, to be retrieved at the clause level and incorporated into a *question* semantic object (Ginzburg and Sag, 2001). Informally, this amounts to analyzing *Combien de livres a lus Paul ?* 'How many books did Paul read?' as asking for the number $n$ such that Paul read $n$ books.

In extraction constructions like (18a), where *combien* is split from the *de*-N′ it licenses, the LOCAL value of *combien* is copied from the common noun's SPR list into SLASH by the second weak head *de* of Figure 3. This SLASH element propagates to the clause level thanks to the standard NON-LOCAL feature principles (SLASH Amalgamation and Inheritance). At that point, *combien* is realized as the filler in a filler-head structure, binding off the element in SLASH (and thus licensing the *de*-N′ phrase in which it originated).[14]

[14] In contrast, in examples like (19b) where the entire NP containing *combien* is extracted, the

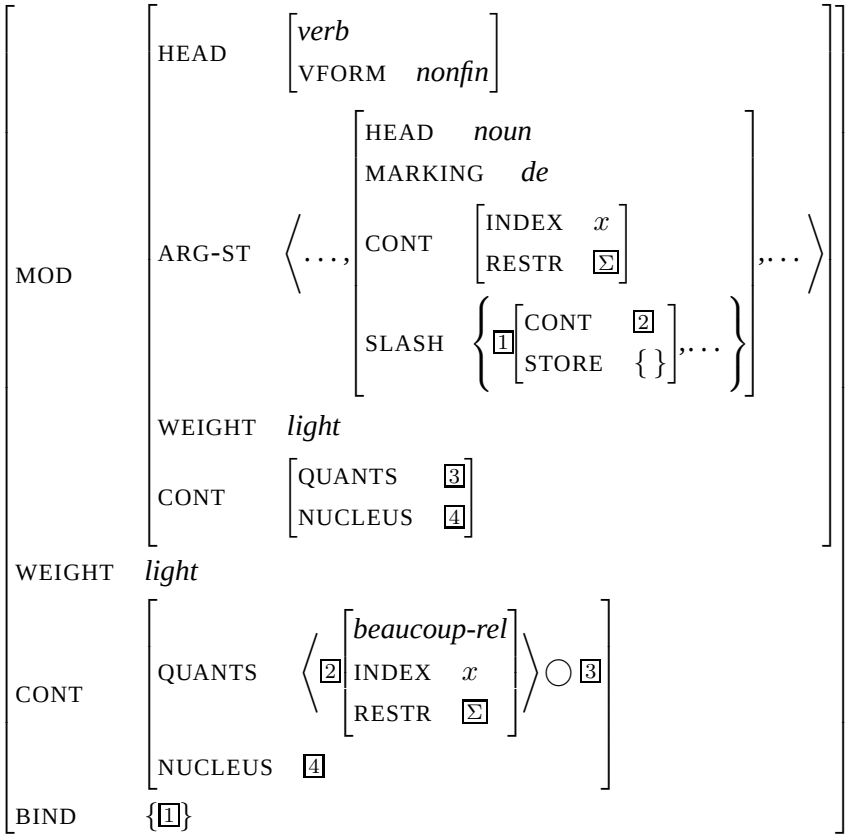


Figure 5: Lexical entry for *beaucoup* (QAD adverb)

### 4.4 QAD (Quantification at a distance)

To account for cases of QAD as in (20), we treat the "floating" adverb as a modifier adjoined to the non-finite verb. The relevant lexical entry for *beaucoup* is given in Figure 5. Via its MOD value, the adverb *beaucoup* selects a verb with a *de*-N′ phrase on its argument structure (whose SLASH set contains a specifier). The specification [WEIGHT *light*] in the MOD value ensures that the modified category is a lexical verb (or a coordination of lexical verbs) and not a branching VP; the same WEIGHT specification on the adverb itself determines its linear position to the left of the modified verb (Abeillé and Godard, 2000).

The adverb binds the SLASH dependency lexically (licensing the autonomous *de*-N′ phrase) and adds its own quantificational semantics to the QUANTS list inherited from the verb; recall that the semantic content shown here is a simplification that does not take into account the data in (23). Because the semantic contribution

weak head *de* of Figure 3 is not involved. The extracted NP is headed by the first weak head of Figure 1, which introduces no SLASH dependency. Such examples are straightforward cases of verbal argument extraction.

does not invoke the STORE mechanism, it is impossible for *beaucoup* to take any wider scope.[15]

### 4.5 Licensing of *de*-N′ in negative contexts

Following de Swart and Sag (2002) and Godard (2004), we assume that French negative words (including the simple negation *pas*) are quantifiers which occur in STORE and are retrieved by a *ne*-marked verb, or by an inherently negative element like the preposition *sans*.[16] Verb words are partitioned into two types, *neg-vb-wd* and *pos-vb-wd*; only *neg-vb-wd* verbs can retrieve the negative quantifiers in their STORE.

We propose that, in addition to retrieving negation, *ne*-marked verbs can also license *de*-N′ phrases (bearing a non-empty SLASH). The combined constraint on *neg-vb-wd* is given here:

(41) *neg-vb-wd*→

$$\left[\begin{array}{ll} \text{CONT}|\text{QUANTS} & \left[\left(\mathit{list}(\mathit{quant\text{-}rel}) \bigcirc \langle \mathit{neg\text{-}quant\text{-}rel}\rangle\right) \oplus \langle \boxed{1},\ldots,\boxed{n}\rangle\right] \\ \text{STORE} & \mathit{list}(\mathit{pos\text{-}quant\text{-}rel} \vee \mathit{param}) \\ \text{BIND} & \left\{\left[\begin{array}{ll} \text{HD}|\text{SPEC}|\text{MARKING} & \mathit{de} \\ \text{CONT} & \boxed{1}\ \mathit{exist\text{-}rel} \\ \text{STORE} & \{\,\} \end{array}\right],\ldots,\left[\begin{array}{ll} \text{HD}|\text{SPEC}|\text{MARKING} & \mathit{de} \\ \text{CONT} & \boxed{n}\ \mathit{exist\text{-}rel} \\ \text{STORE} & \{\,\} \end{array}\right]\right\} \end{array}\right]$$

The *neg-vb-wd* retrieves all stored negative quantifiers (there must be at least one). The verb can also bind specifiers that have been introduced into SLASH by *de*-N′ phrases; these elements can be identified by the feature [HEAD|SPEC|MARKING *de*]. For each of these specifiers, the verb adds an existential quantifier scoping over the corresponding *de*-N′ phrase to the end of its QUANTS list. (The weak head entry in Figure 3 ensures that the INDEX and RESTR of the specifier in SLASH are identified with those of the specified N′.) These existential quantifiers are scoped below the ordinary quantifiers retrieved from STORE, in accordance with the observations in (24–25).[17]

[15] We assume that adjuncts can modify both the NUCLEUS and QUANTS of the verb. This is independently necessary if we are to account for modal, habitual, and frequency adverbs, which can all outscope quantifiers. We must also assume that non-local features (SLASH, STORE) are amalgamated by left-adjoined adverbs and pass from non-head daughter to mother in these head-adjunct phrases.

[16] Here we focus on retrieval of negation by verbs. See Godard (2004) for a fuller discussion of possible negation retrieval sites in French.

[17] The constraint in (41) states that specifiers introduced by *de*-N′ phrases are the only SLASH elements that can be bound lexically; a more elaborate constraint would be needed to make the present analysis compatible with the treatment of *en*-cliticization out of NPs proposed in Miller and Sag (1997).

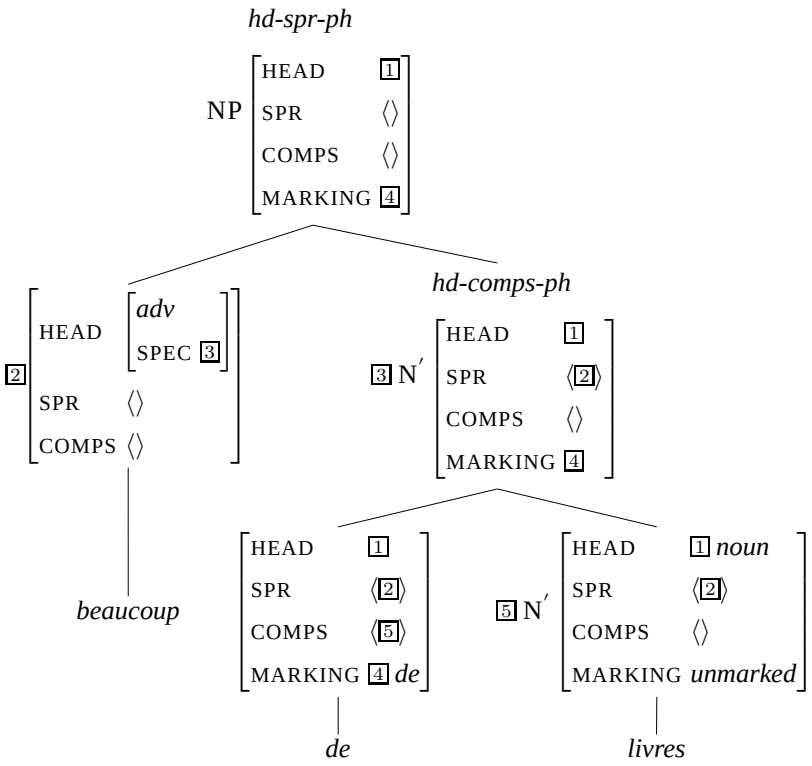


Figure 6: Local quantification [Deg *de* N′] (see section 4.1)

# 5 Final remarks

The proposals presented here rely heavily on the notion of weak head, an alternative to the category *marker* of standard HPSG (Tseng, 2002).[18] French *de* cannot be analyzed as a marker, because it has to be able to introduce its own valence requirements—recall the lexical entry in Figure 3—and semantics, in the case of partitive *de* as in (8), for instance.

For similar reasons, the recent proposals of Van Eynde (2004) cannot be applied directly to *de*: his "functor" elements primarily contribute a new MARKING value, much like standard HPSG markers. Unlike markers, functors can make a semantic contribution, but they still cannot modify the valence of their sister category. Furthermore, Van Eynde's treatment of specifiers as functors is also incompatible with the crucial idea in our analysis of *de*-N′ licensing: specifiers must be extractable arguments.

Van Eynde's notion of "minor category" could be useful for capturing the functional restrictions that characterize *de*—e.g., it cannot be modified by or conjoined

[18]The weak head approach has also been used in the analysis of coordinating conjunctions (Abeillé, 2003).

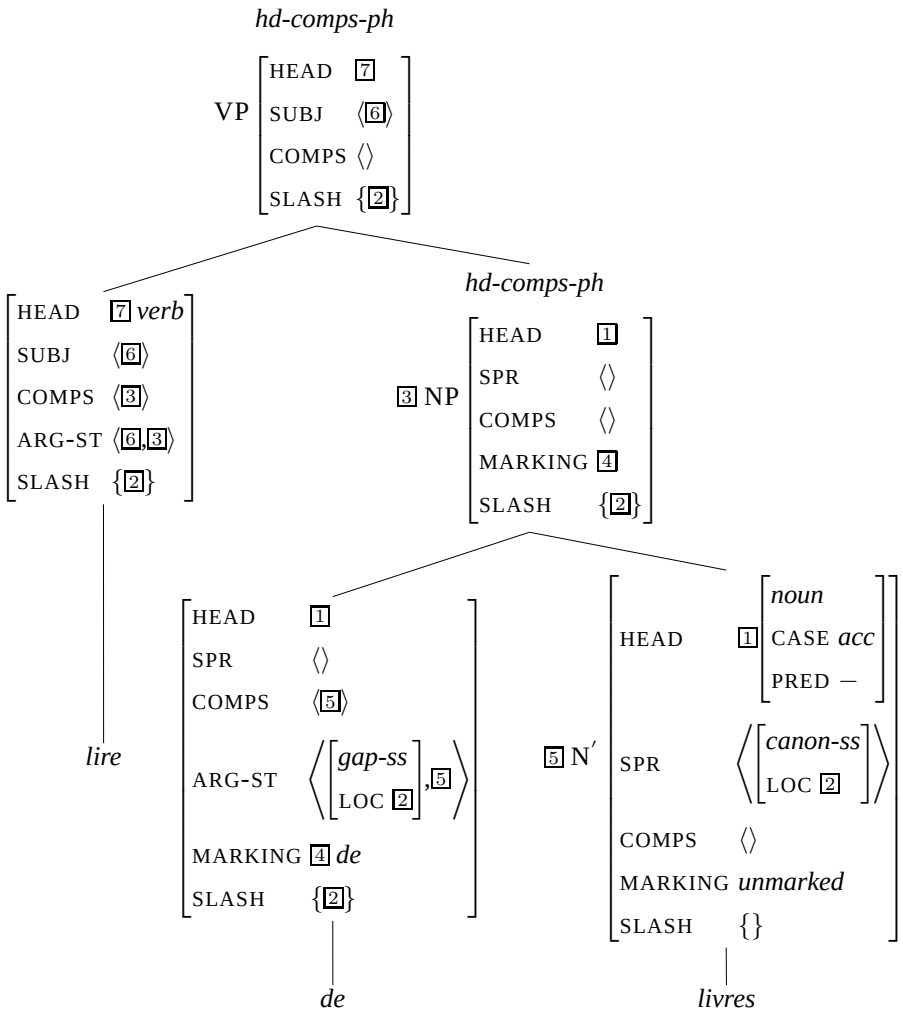


Figure 7: Autonomous *de*-N′ in need of licenser (see section 4.2)

with another element, or used in isolation. But these restrictions apply to all uses of *de*, including oblique (i.e., prepositional) uses that would clearly be "major" in Van Eynde's system. It is possible that the details of the major/minor dichotomy could be adapted; alternatively these properties of *de* could be analyzed as part of the theory of syntactic weight (Abeillé and Godard, 2000).

## References


Abeillé, Anne. 1997. Fonction ou position objet ? Deuxième partie. *Le gré des langues* 12, 8–33.

Abeillé, Anne. 2003. A lexicalist and construction-based approach to coordination. In Stefan Müller (ed.), *Proceedings of the 10th International HPSG Conference*, pages 5–25, Stanford, CA: CSLI Publications.

Abeillé, Anne and Godard, Danièle. 1999. A Lexical Approach to Quantifier Floating in French. In Gert Webelhuth, Jean-Pierre Koenig and Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, pages 81–96, Stanford, CA: CSLI Publications.

Abeillé, Anne and Godard, Danièle. 2000. French Word Order and Lexical Weight. In Robert D. Borsley (ed.), *The Nature and Function of Syntactic Categories*, volume 32 of *Syntax and Semantics*, pages 325–360, San Diego: Academic Press.

Abeillé, Anne and Godard, Danièle. 2003. The Syntactic Flexibility of Adverbs: French Degree Adverbs. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on HPSG*, pages 26–46, Stanford, CA: CSLI Publications.

Abeillé, Anne, Godard, Danièle, Bonami, Olivier and Tseng, Jesse. 2003. The syntax of French *à* and *de*: An HPSG analysis. In Patrick Saint-Dizier (ed.), *Proceedings of the ACL-SIGSEM Workshop on the linguistic dimensions of prepositions and their use in computational linguistics formalisms and applications*, pages 133–144, Toulouse: IRIT.

Bonami, Olivier, Godard, Danièle and Marandin, Jean-Marie. 1999. Constituency and word order in French subject inversion. In Gosse Bouma, Erhard Hinrichs, Geert-Jan Kruijff and Richard T. Oehrle (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, pages 21–40, Stanford, CA: CSLI Publications.

Borsley, Robert D. 2004. An approach to English comparative correlatives, paper presented at the 11th International HPSG Conference, Leuven, August 2004.

de Swart, Henriëtte and Sag, Ivan A. 2002. Negation and negative concord in Romance. *Linguistics and Philosophy* 25, 373–417.

Doetjes, Jenny. 1997. *Quantifiers and Selection. On the distribution of quantifying expressions in French, Dutch and English*. Ph. D.thesis, Holland Institute of Generative Linguistics, Leiden University, published by Holland Academic Graphics.

Gaatone, David. 1971. *Etude descriptive du système de la négation*. Geneva: Droz.

Gaatone, David. 1992. *De* négatif entre syntaxe et sémantique ; Réflexions sur quelques propriétés du déterminant *de*. *Langue française* 94, 93–102.

Ginzburg, Jonathan and Sag, Ivan A. 2001. *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford, CA: CSLI Publications.

Godard, Danièle. 2004. French Negative Dependency. In Francis Corblin and Henriëtte de Swart (eds.), *Handbook of French Semantics*, Stanford, CA: CSLI Publications.

Gross, Maurice. 1967. Sur une règle de cacophonie. *Langages* 7, 105–119.

Huot, Hélène. 1981. *Constructions infinitives du français : le subordonnant* de. Geneva: Droz.

Kayne, Richard S. 1981. ECP Extensions. *Linguistic Inquiry* 12, 93–133.

Kehler, Andrew. 2002. *Coherence, Reference, and the Theory of Grammar*. CSLI Publications.

Kupferman, Lucien. 2004. *Domaines prépositionnels et domaines quantificationnels : Le mot* de. Paris: Duculot.

Miller, Philip H. and Sag, Ivan A. 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15, 573–639.

Milner, Jean-Claude. 1978. *De la syntaxe à l'interprétation*. Paris: Le Seuil.

Muller, Claude. 1997. *De* partitif et la négation. In Danielle Forget, Paul Hirschbühler, France Martineau and María-Luisa Rivero (eds.), *Negation and Polarity*, pages 251–270, John Benjamins.

Obenauer, Hans-Georg. 1994. Aspects de la syntaxe A-barre. Effets d'intervention et mouvements des quantifieurs. Thèse d'état, Université Paris 8.

Tseng, Jesse. 2002. Remarks on Marking. In Frank van Eynde, Lars Hellan and Dorothee Beermann (eds.), *Proceedings of the 8th International HPSG Conference*, pages 267–283, Stanford, CA: CSLI Publications.

Van Eynde, Frank. 2004. Minor Adpositions in Dutch. *Journal of Comparative Germanic Linguistics* 7, 1–58.