

The syntax of French *de*-N' phrases

Anne Abeillé

Olivier Bonami

Danièle Godard

Jesse Tseng

Univ. Paris 7/LLF

Univ. Rennes 2/LLF

CNRS, LLF/Univ. Paris 7

CNRS, Loria

E-mail: {abeille|dgodard}@linguist.jussieu.fr, olivier.bonami@uhb.fr, tseng@loria.fr

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 requires a different treatment. After outlining a general distinction between “oblique” and “non-oblique” uses of *de*, we provide a detailed account of constructions where *de* combines with an N'. We provide a unitary analysis of *de* in three constructions (*combien* extraction, “quantification at a distance”, and negative contexts) which have been treated separately in previous accounts.

1 The dual syntactic nature of *de*

The uses of *de* can be partitioned into two classes, based on a number of syntactic criteria. “Oblique” *de* disallows extraction out of the phrase it combines with (1a), and can combine with a coordination of phrases (1b). Oblique *de*-phrases can appear in complement or adjunct positions, but not in subject position (1c). Using these tests, we find oblique *de* in combination with NP and N' as in (1), but also with PP and AP/AdvP (2).

- (1) a. *un film dont je me souviens [de la fin]
'a film of which I remember the ending'
b. J'ai besoin de [cette farine et cette levure].
'I need this flour and this yeast.'
c. *[De mort] est la seule façon efficace de menacer ces gens.
'With death is the only effective way to threaten these people.'
- (2) a. Il surgit [de derrière le rideau].
'He jumps out from behind the curtain.'
b. quelque chose [de plus traditionnel]
'something more traditional'

In “non-oblique” uses, *de* behaves very differently: it allows extraction out of its sister phrase (3a), it cannot take wide scope over a coordination of phrases (3b), and non-oblique *de*-phrases can occur in subject position (3c). In addition to these examples with N' and infinitival VP, non-oblique *de* also combines with definite NPs in so-called “partitive” constructions (4).

- (3) a. un auteur dont je n'ai pas lu [de livres]
'an author who I haven't read any books by'
b. *On nous a apporté plein [de pain et vin].
'They brought us loads of bread and wine.'
c. [De sortir un peu] te ferait du bien.
'Getting out a bit would do you some good.'

HEAD	prep						
MARKING	<i>de</i>						
COMPS	<table border="1"> <tr> <td>HEAD</td> <td>¬ verb</td> </tr> <tr> <td>MARKING</td> <td>¬ <i>de</i></td> </tr> <tr> <td>COMPS</td> <td>⟨ ⟩</td> </tr> </table>	HEAD	¬ verb	MARKING	¬ <i>de</i>	COMPS	⟨ ⟩
HEAD	¬ verb						
MARKING	¬ <i>de</i>						
COMPS	⟨ ⟩						
SLASH	{ }						

Figure 1: oblique *de* as a preposition

HEAD	①												
MARKING	<i>de</i>												
SUBJ	②												
SPR	③												
COMPS	<table border="1"> <tr> <td>HEAD</td> <td>①</td> </tr> <tr> <td>MARKING</td> <td>unmarked</td> </tr> <tr> <td>SUBJ</td> <td>②</td> </tr> <tr> <td>SPR</td> <td>③</td> </tr> <tr> <td>COMPS</td> <td>⟨ ⟩</td> </tr> <tr> <td>CONT</td> <td>④</td> </tr> </table>	HEAD	①	MARKING	unmarked	SUBJ	②	SPR	③	COMPS	⟨ ⟩	CONT	④
HEAD	①												
MARKING	unmarked												
SUBJ	②												
SPR	③												
COMPS	⟨ ⟩												
CONT	④												
CONT	④												

Figure 2: non-oblique *de* as a weak head

- (4) un courrier contenant [de la poudre blanche suspecte]
'a letter containing suspicious white powder'

The properties of oblique *de*-phrases are straightforwardly accounted for by analyzing *de* as an ordinary preposition, satisfying the lexical description in Fig. 1. French PPs are extraction islands, and cannot be used as subjects. Non-oblique *de*-phrases, on the other hand, behave unlike PPs, and more like NPs or VPs. We propose that non-oblique *de* is a “weak head”—that is, a word that shares its HEAD value with its complement. Fig. 2 shows one lexical entry for the weak head *de*, used with nominal and verbal complements. For example, to form a partitive phrase as in (4), *de* selects an NP complement (starting with a definite, demonstrative or possessive specifier), and introduces partitive semantics. The resulting *de*-NP has the HEAD value *noun* and has the grammatical properties of an ordinary NP with respect to distribution, extraction, etc.

Oblique and non-oblique *de* do share one crucial property: the MARKING value *de*, which then propagates to their projections. This explains why all pronominalizable *de*-phrases alternate with the clitic *en*, despite their otherwise highly divergent gram-

mathematical properties. The principles governing *encliticization* refer only to the feature [MARKING *de*]. The MARKING attribute is also used to prevent stacked occurrences of *de*: *de* never combines with a [MARKING *de*] complement.¹

Finally, we note that the partition into prepositional and weak head 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 turn to (non-oblique) nominal phrases of the form *de-N'*, which have restricted distribution and must always be licensed by other material. In the simplest case, they are licensed locally by a degree expression from a class including adverbs (*beaucoup*, *infiniment*), nouns (*nombre*, *quantité*), or the invariable form *plein*. The resulting NPs (e.g., *beaucoup de livres* ‘a lot of books’) have no distributional restrictions; they can appear as subject or complement of a verb, or complement of a preposition. 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 clearly relate this type of uses to modifier uses of the same items (Abeillé & Godard, 2003);² here we treat degree adverbs in [Adv *de N'*] phrases as atomic binary quantifiers, for the sake of brevity.

There are two variants of *de-N'* licensed by degree quantification. First, the degree expression can be the adverbial *wh*-word *combien* ‘how much/many’, extracted and realized in sentence-initial position (5). Or alternatively, a small number of degree adverbs (*beaucoup*, *trop*, *assez*) can “float” immediately to the left of an infinitive or past participle (6a). In such cases of

¹*De* also displays some morphophonological idiosyncrasies: it always undergoes vowel elision conditioned by the following context, and contraction with the specifier forms *le* and *les*, giving rise to *du* and *des*. Again, these phenomena are not conditioned by the oblique vs. non-oblique status of *de*.

²Abeillé and Godard (2003) attempt to relate the two uses by treating degree adverbs uniformly as modifiers, even in the [Adv *de N'*] construction—thus incorrectly predicting that phrases with this construction should lack quantificational force. A more adequate way to relate the two 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 modifier use, and to specify the size of the group which is also quantified over in the [Adv *de N'*] construction.

- (i) Paul a beaucoup dormi. ‘Paul slept a lot.’
 $\exists e[\text{sleep}(e, p) \wedge \text{a-lot}(e, S)]$
- (ii) Beaucoup d’enfants dorment. ‘Many children are sleeping.’
 $\exists e \exists X [\text{children}(X) \wedge \text{sleep}(e, X) \wedge \text{a-lot}(X, S)]$

“quantification at a distance” (QAD), the degree adverb always takes narrow scope (7).³

- (5) Combien as-tu lu [de livres en latin] ?
 ‘How many have you read DE books in Latin?’
- (6) a. Il va beaucoup lire [de livres] / Il a beaucoup lu [de livres].
 ‘He will MANY read DE books / He has MANY read DE books’
- b. Jean a vu / apprécié [beaucoup de films].
 ‘Jean has seen/appreciated many films’
- c. Jean a beaucoup vu / ??apprécié [de films].
 ‘Jean has MANY seen/appreciated DE films’
- (7) a. Paul recevra chaque étudiant qui a lu [beaucoup de livres].
 ‘Paul will meet every student who read a large number of books’ or
 ‘There are a large number of books *x* such that P. will meet every student who read *x*’
- b. Paul recevra chaque étudiant qui a beaucoup lu [de livres]. (first reading only)

Finally, *de-N'* phrases can be licensed by negation. In such cases the *de-N'* is interpreted as an existential quantifier in the immediate scope of negation; thus (8a) and (8b) share a reading similar to (8c), but only (8b) can mean (8d).⁴

- (8) a. Paul n’a pas lu de livre.
 b. Paul n’a pas lu un livre.
 c. Paul did not read any book.
 d. There is a book Paul did not read.

3 Autonomous *de-N'*s are extraction sites

The cases of *combien* extraction and *de-N'* in negative contexts have not traditionally been considered to be related. Intriguingly, however, there are at least three properties that unite them.

Distribution in both cases the *de-N'* has the same distribution: it can be a direct object (5b–8a) or an inverted subject (9a,10a), but not a preverbal subject (9b,10b) or the complement of a preposition (9c,10c).

- (9) a. Combien dis-tu que sont venus [de clients] ?
 ‘You say that how many clients came?’

³In the QAD construction, there is some form of semantic selection of the verb by the degree adverb (see the examples in (6b,c)). This has been analyzed as showing that the adverb quantifies over occasions rather than individuals (Obenauer, 1994; Doetjes, 1997). However the data are quite tricky, and it is not completely clear 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.

⁴Note that *de-N'* phrases in negative contexts are not (ordinary) NPIs, since they are only licensed by explicit negative words. They also differ from negative words in that they never have negative force on their own.

- b. *Combien dis-tu que [de clients] sont venus?
 c. *Combien as-tu voté contre [de projets] ?
 ‘How many projects did you vote against?’
- (10) a. un endroit où ne vont pas [d’enfants]
 ‘a place where no children go.’
 b. *un endroit où [d’enfants] ne vont pas
 c. *Paul n’a jamais voté contre [de projet].
 ‘Paul never voted against any project.’

Long distance Just like the extraction construction with *combien*, the dependency between the negative word and *de-N'* is a long distance dependency.

- (11) 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.’

Island constraints Quite unexpectedly, the dependency between the negative word and the *de-N'* obeys extraction island constraints. The *de-N'* can be embedded in a complement clause (11a), but it cannot be embedded in a relative clause or in a subject clause (12a-b). Coordinations of constituents containing *de-N'*s are possible, but it is not possible to coordinate two constituents if only one of them contains a *de-N'* (13). Finally, the *de-N'* cannot be embedded in a PP (10c).

- (12) a. *Que Paul ait lu [de livre] ne m’a pas surpris.
 ‘That Paul read a book did not surprise me.’
 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.’
- (13) a. 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 or to watch a movie with him in it.’
 b. *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 or to go the movies tonight.’

Discussion Given these properties, we propose that *de-N'* is at the bottom of a SLASH dependency in negative contexts. This dependency terminates at the node where the negation is retrieved. The QAD construction, on the other hand, does not share exactly the same properties as *combien* extraction and *de-N'* in negative contexts. But QAD does not violate any of the constraints observed—i.e., it is possible in a subset of the cases where the other two constructions are possible. Thus we may assume that QAD rests on the same basic syntax for the *de-N'* phrase, but that the construction puts further restrictions on its distribution.

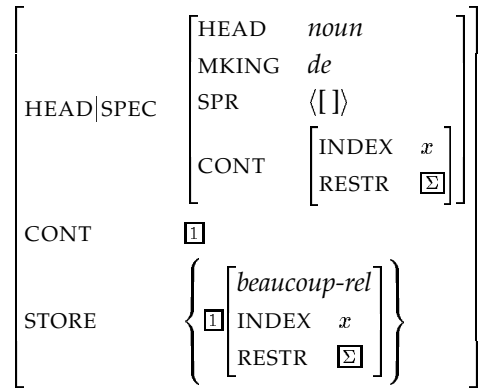


Figure 3: *beaucoup* in [Adv *de N'*] contexts

4 An HPSG analysis of *de-N'*

We assume throughout 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: [Adv *de N'*]

We analyze [Adv *de N'*] as an ordinary head-specifier phrase, where the specifier is quantificational. Fig. 3 shows the relevant lexical entry for *beaucoup*. It is a typical specifier entry, except that *beaucoup* selects a [MARKING *de*] head. Accordingly *beaucoup* can combine only with *N'*s which are headed by the weak head *de*.

4.2 No unlicensed *de-N'*

To account for the distribution of *de-N'* we assume two distinct lexical entries for the weak head *de*. The first, shown above in Fig. 2, is also used for *de*-VPs. In this entry both subject and specifier are raised. We assume independently that the first argument of a noun (i.e., its specifier) must be canonical (14). As a consequence, any *de-N'* headed by this weak head *de* will have a canonical specifier on SPR. But verbs do not allow SPR-unsaturated direct arguments unless they are predicative (15); thus in the absence of a specifier like *beaucoup*, *de-N'*s headed by the weak head in Fig. 2 cannot occur as arguments of verbs.⁵

$$(14) \text{cn-wd} \rightarrow [\text{ARG-ST } \langle \text{canon-ss} \rangle \oplus \text{list}(\text{synsem})]$$

$$(15) \text{vb-wd} \rightarrow [\text{ARG-ST } \text{list}([\text{PRED } +] \vee [\text{SPR } \langle \rangle])]]$$

The second entry for *de*, given in Fig. 4, is specific to *de-N'*s introducing a SLASH dependency. This head

⁵As it stands, the entry in Fig. 2 allows *de* to combine with full NPs, predicting sentences like *Paul voit de Marie / de plusieurs poissons ('Paul sees DE Marie / DE several fish'). This can be excluded by further stating that the weak head must inherit either a subject or a specifier.

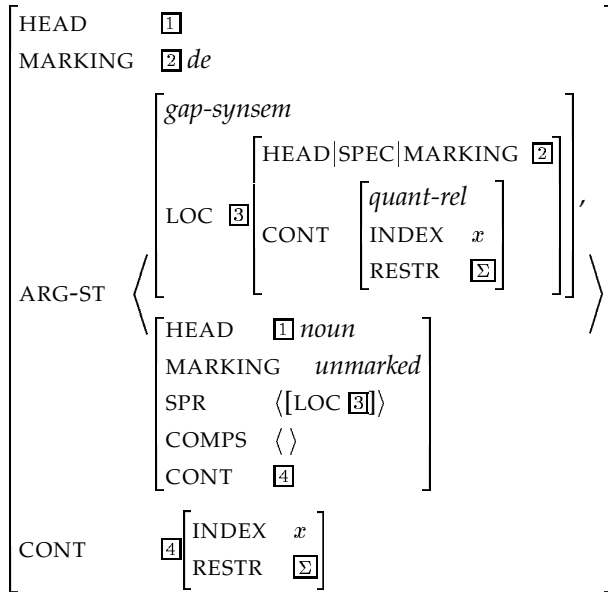


Figure 4: a second weak head *de*

takes a nominal complement unsaturated for SPR, and has a *gap* first argument which shares the LOCAL value of the complement’s SPR.⁶ We assume that *gap* elements on ARG-ST are not mapped to the SPR list; hence the weak head of Fig. 4 gives rise to saturated *de-N*’s, which are allowed as arguments of verbs (see (15)), but their SLASH contains a *local* object identified with the noun’s unrealized specifier. That these *de-N*’s can only occur as complements or inverted subjects is a consequence of the LP rule (16).

$$(16) \quad V \prec \text{NP}_x \left[\begin{array}{l} \text{MARKING } de \\ \text{SLASH } \left\{ \left[\text{CONT } \left[\begin{array}{l} \text{quant-rel} \\ \text{IND } x \end{array} \right] \right] \dots \right\} \end{array} \right]$$

4.3 Extracted *combien*

The lexical entry for *combien* is given in Fig. 5. We propose that *combien* is a quantifier with an extra argument for a quantity parameter; this parameter is put on STORE to be retrieved at clause level and enter into the formation of a *question* semantic object (Ginzburg & 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. Given the description in Fig. 5, the possibility of extracting *combien* falls out directly from the general treatment of *wh*-interrogatives.

4.4 QAD

As shown in Fig. 6, we capture the distribution of QAD by assuming that a QAD adverb modifies a verb

⁶Note that although the weak head selects a common noun complement and inherits its HEAD value, *de* itself is not of type *cn-wd*, and thus does not violate (14).

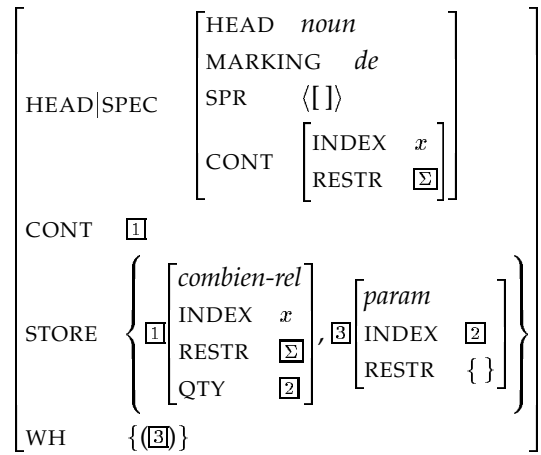


Figure 5: lexical entry for *combien*

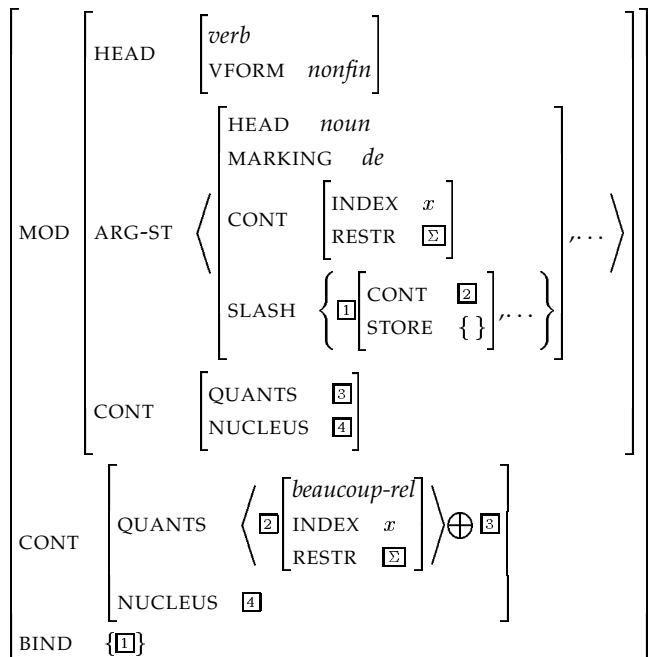


Figure 6: *beaucoup* in QAD contexts

with a *de-N*’ on its argument structure. This *de-N*’ is slashed for a specifier; the adverb binds the SLASH dependency lexically, and adds the specifier’s semantic contribution to the verb’s QUANTS list. Since the quantifier is not put in STORE, it will never get a chance to scope higher.⁷

4.5 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

⁷Note that we assume that adjuncts may 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 also assume that non-local features (SLASH, STORE) are amalgamated by adjuncts and pass from non-head daughter to mother in such head-adjunct phrases.

STORE and are scoped on a *ne* marked verb.⁸ Thus the general treatment of negation requires that verb words are partitioned into two types, *neg-vb-wd* and *pos-vb-wd*, where only *neg-vb-wd* verbs retrieve all negative quantifiers in their STORE.

The distribution and interpretation of *de-N'*s in negative contexts can be accounted for by elaborating the constraint on *neg-vb-wd*. Informally, we require these verbs to also scope an existential quantifier for each of the *de-N'*s in their dependency. Since these correspond to identifiable elements on SLASH, there is no need to refer directly to *de* phrases: any *local* object in SLASH with the feature [HEAD|SPEC|MARKING *de*] must be scoped. More precisely, the constraint in (17) states that all elements in BIND must be [HEAD|SPEC|MARKING *de*] and contribute an existential quantifier as their CONTENT. The weak head entry in Fig. 4 ensures that the INDEX and RESTR of this quantifier will be identified with those of the noun in a *de-N'* phrase. These existential quantifiers are then scoped below the ordinary quantifiers retrieved from ARG-ST.⁹

(17) *neg-vb-wd* →

$$\left[\begin{array}{l} \text{ARG-ST} \left\langle \left[\text{STORE } \Sigma_1 \right], \dots, \left[\text{STORE } \Sigma_n \right] \right\rangle \\ \text{CONT|QUANTS} \left[\oplus \left\langle \left[\Sigma_1 \cup \dots \cup \Sigma_n \right] \setminus \Sigma \right\rangle \right] \\ \text{STORE } \Sigma \left(\text{list}(\text{pos-quant-rel} \vee \text{param}) \right) \\ \text{BIND} \left\{ \begin{array}{l} \left[\begin{array}{l} \text{HD|SPEC|MARKING } de \\ \text{CONT } \Sigma \text{ exist-rel} \\ \text{STORE } \{ \} \end{array} \right], \dots, \\ \left[\begin{array}{l} \text{HD|SPEC|MARKING } de \\ \text{CONT } \Sigma \text{ exist-rel} \\ \text{STORE } \{ \} \end{array} \right] \end{array} \right\} \end{array} \right]$$

5 Final remarks

Our proposals rely heavily on the notion of weak head, an alternative to the category *marker* of standard HPSG (Tseng, 2002). French *de* cannot be analyzed as a marker, because it has to be able to introduce its own valence requirements (as in Fig. 4) and semantics (in the case of partitive *de*).

⁸This is a simplification, since the inherently negative preposition *sans*, as well as some adjectives, can also be negation retrieval sites (Godard, 2004). For the sake of brevity we concentrate on cases where the negation retrieval site is a verb.

⁹A number of details are skipped here. Further constraints are necessary to ensure that (i) *neg-vb-wds* scope at least one *neg-quant-rel*, and (ii) *pos-vb-wds* can scope only *pos-quant-rels*. A further simplification is that according to the constraint in (17), specifiers of *de-N'*s 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).

For the same reasons, the recent proposals of Van Eynde (2004) cannot be applied directly to *de*: just as for standard marker daughters, the only grammatical information contributed by his “functor” elements is a new MARKING value. His treatment of specifiers as functors is also incompatible with the crucial idea in our analysis of *de-N'*: specifiers must be extractable arguments.

Van Eynde’s notion of “minor category” might be useful for capturing the functional restrictions that characterize *de*—e.g., it cannot be modified by or conjoined 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 & Godard, 2000).

References

- Abeillé, A., & Godard, D. (2000). French word order and lexical weight. In R. D. Borsley (Ed.), *The nature and function of syntactic categories* (Vol. 32, pp. 325–360). San Diego: Academic Press.
- Abeillé, A., & Godard, D. (2003). The syntactic flexibility of adverbs: French degree adverbs. In S. Müller (Ed.), *Proceedings of the 10th international conference on HPSG* (pp. 26–46). Stanford, CA: CSLI Publications.
- de Swart, H., & Sag, I. A. (2002). Negation and negative concord in Romance. *Linguistics and Philosophy*, 25, 373–417.
- Doetjes, J. (1997). *Quantifiers and selection. On the distribution of quantifying expressions in French, Dutch and English*. PhD Thesis, Holland Institute of Generative Linguistics, Leiden University. (Published by Holland Academic Graphics.)
- Ginzburg, J., & Sag, I. A. (2001). *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.
- Godard, D. (2004). French negative dependency. In F. Corblin & H. de Swart (Eds.), *Handbook of French semantics*. Stanford, CA: CSLI Publications.
- Miller, P. H., & Sag, I. A. (1997). French clitic movement without clitics or movement. *Natural Language and Linguistic Theory*, 15, 573–639.
- Obenauer, H. (1994). *Aspects de la syntaxe A-barre. Effets d'intervention et mouvements des quantifieurs*. (Thèse d'état, Université Paris 8)
- Tseng, J. (2002). Remarks on marking. In F. van Eynde, L. Hellan, & D. Beermann (Eds.), *Proceedings of the 8th International HPSG Conference* (pp. 267–283). Stanford, CA: CSLI Publications.
- Van Eynde, F. (2004). Minor Adpositions in Dutch. *Journal of Comparative Germanic Linguistics*, 7, 1–58.