French Reportive *Comme* Clauses: a case of parenthetical adjunction

Marianne Desmets
Paris X – Nanterre University
Paris 7 – Denis Diderot University

Laurent Roussarie
Paris 7 – Denis Diderot University
ENS Fontenay/Saint-Cloud

1 Introduction[1]

The French system of reported speech exhibits a particular construction introduced by the adverb *comme*—often translated by *as* in English. We call it Reportive *Comme* Clause (henceforth RCC)[2]. This construction has not received much attention in the French literature, mostly because the range of uses of *comme* is wide and fairly complex. But RCCs combine a large number of interesting syntactic, lexical and semantic features that justify considering them as a distinct construction[3]. Among works on the English counterpart of RCCs, we can mention Ross (1967), Partee (1973), and more recently Lapointe (1991) which presents a null operator analysis.

We propose an analysis of this French construction as a parenthetical adjunct clause. We show that RCCs are extraction contexts and a subtype of free relative clauses. Contra Lapointe, we claim that it is not necessary to deal with any null operators or to posit any empty categories in order to account for RCC extraction. Instead, following current HPSG accounts, we propose a SLASH treatment for this filler-gap phrase. We integrate RCCs in the type hierarchy of phrases (proposed by Sag 1997, Abeillé *et al.* 1998 for French) as a subtype of head-adjunct phrase and head-filler phrase. We make a distinction between parenthetical adjuncts and head-modifying adjuncts that allows a simplified treatment for parentheticals. We also present an account of direct speech and quoted argument selection which involves a new type of non-canonical realization.

First, we present the syntactic properties of RCCs. Then, we move to the semantic specifications of the construction. In the last part of the paper, we present a synthesis of our proposals and an HPSG formalization of the analysis.

[1] We would like to thank anonymous reviewers for valuable comments and I. Sag, O. Bonami, D. Godard, A. Delaveau, J. Sadock, E. Hinrichs, E. Bender, J. Tseng, A. Abeillé, for important contributions, discussions and comments.

[2] The term 'reportive' is borrowed from Lapointe (1991).

[3] The present work is based upon an on going research work (Desmets (2001), Doctorat thesis).

Proceedings of the 7th International HPSG Conference, UC Berkeley (22-23 July, 2000).
Dan Flickinger and Andreas Kathol (Editors).
2001. CSLI Publications. http://csli-publications.stanford.edu/

2 Syntactic properties of RCCs

2.1 An obligatory anaphoric relation

RCCs exhibit an obligatory anaphoric relation between the object argument of the reportive verb and the element in the scope of the *comme*-clause. In most cases, this element is the whole main clause (as in 1a), but it may be only a part of it, in the case of a quotative use (as in 1b).

(1) a. La situation est critique, comme l'a affirmé hier le premier ministre.
'the situation is critical, as it-claimed yesterday the prime minister'
The situation is critical, as the Prime Minister claimed.

b. La situation mérite d'être traitée de façon "politiquement correcte", comme disent les Américains.
'the situation deserves to be treated in-a-way "politically correct", as say the American-PLUR'
The situation deserves to be treated in a "politically correct" way, as American people say.

The object argument of the reportive verb is never canonically realized, as can be seen in (2). But, in contrast to the corresponding English construction, which prohibits the realization of the object, RCCs allow some alternation in the realization of the object. It can be a pronominal affix unmarked for agreement (ex. 3a)—accusative *le* or dative *y*, depending on the verb subcategorization requirements—or it can be a gap (ex. 3b):

(2) *La bourse s'est effondrée, comme le New York Times avait prévu *cela*/ *cet événement*.
'the Stock Market it-be crashed, as the NY Times predicted that/ this event'

(3) a. La bourse s'est effondrée, comme *l'*avait prévu le New York Times.
'the Stock Market it-be crashed, as it-predicted The NY Times'
The Stock Market crashed, as the NY Times predicted.

b. Les pluies gagneront la côte Ouest ce soir, comme a dit_ le journaliste.
'the rain reach-FUT the coast west this evening, as said the journalist'
Rain will reach the west coast this evening, as the journalist said.

We will demonstrate that this obligatory anaphoric relation is a constructional constraint.

First, we observe that the non-canonical realization of the object is crucial to obtain the proper reported speech semantics. A canonical realization of the object is not impossible *per se*, but when the object is realized, it clearly changes the interpretation of the *comme*-clause into a manner modifier adjunct (the same way it changes in English Reportive *as* clauses):

(4) La bourse s'est effondrée, comme le New York Times avait prévu qu'elle s'effondrerait.
'the Stock Market it-be crashed, as the NY Times predicted that it crash-COND'
The Stock Market crashed, as the NY Times predicted it would.

The interpretation obtained in (4) is no longer 'the Stock Market crashed, which the NY Times predicted', but 'the Stock Market crashed the way that the NY Times predicted it would'. Hence, the *comme* adjunct in (4) is construed as a predicate modifier.

Second, in the following data, we observe that RCCs show clear extraction properties. This means that the lower verb of the reportive clause (a verb of reported speech that we will call the "reportive" verb) has a non-empty SLASH value. Now, the source of the object non-canonicity needs to be established. There are two hypotheses here: (1), the object non-canonicity results from a syntactic constraint, where the object argument of the reportive verb is the triggering element of the syntactic dependency. Or (2) the obligatory anaphoric relation between the main clause and the object argument of the reportive verb stems from a semantic constraint, involving a particular lexical realization of the object. That would imply that it is not the object argument that is extracted, but some other dependent of the reportive verb.

Our claim is that only the second proposal is a correct analysis for RCCs. First, we will prove that the behavior of the object argument does not fit with French unbounded dependency regularities, which rules out hypothesis (1). Then, we present arguments in favor of hypothesis (2). The analysis we will propose involves an adverbial extraction.

2.2 Extraction context properties

We observe that RCCs are unbounded dependency structures. First, they may contain long-distance dependencies, as in (5):

(5) a. L'enfer, c'est les autres, comme je crois qu'a dit_ Sartre.
'Hell is other people, as I think that said Sartre'
Hell is other people, as I think that Sartre said.

b. La bombe explosera à 3h précises, comme le témoin pense que *l*'a dit le terroriste, bien qu'il n'en soit plus très sûr.
'the bomb explode-FUT at 3:00 exactly, as the witness thinks that it-said the terrorist, although he not be-SUBJUNCT of-it so sure'
The bomb will explode exactly at 3:00, as the witness thinks that the terrorist said, although he is not so sure.

Second, subject inversion frequently arises. This phenomenon is known as a criterion for extraction contexts in French, as discussed in Hukari and Levine (1995), Abeillé *et al.* (1998)—henceforth AGS98. It is what we call an extraction-triggered subject inversion (ETI)[4], after Bonami *et al.* (1998) :

(6) a. La bombe explosera à 3h précises, comme le témoin pense qu'a dit_ le terroriste, bien qu'il n'en soit plus très sûr.

b. La bombe explosera à 3h précises, comme le témoin pense que l'a dit le terroriste, bien qu'il n'en soit plus très sûr.

Third, they are sensitive to Island constraints (Ross 1967; Godard 1988). The reportive verb cannot be realized in an embedded relative clause (CNPC) ex.7), in an interrogative embedded clause (ex.8), or in a cleft clause (ex.9):

(7) a. *Comme l'élève se souvient de l'écrivain qui a écrit_, l'enfer, c'est les autres.
'As the student remembers the writer who wrote, hell is other people'.

b. *Comme l'élève se souvient de l'écrivain qui *l*'a écrit, l'enfer, c'est les autres.
'As the student remembers the writer who it-wrote, hell is other people'

[4] ETI is a more precise term than the traditional 'Stylistic Inversion' (Kayne (1973) and Kayne and Pollock (1978)).

(8) a. *Comme Jean demandait qui a dit_, l'enfer, c'est les autres.
'As John was asking who wrote, hell is other people'

b. *Comme Jean demandait qui *l'*a dit, l'enfer, c'est les autres.
'As John was asking who it-wrote, hell is other people'

(9) a. *Comme c'est le rapport des experts qui nous conduit à dire_, les chiffres ont été truqués.
'As it is the report of experts that leads us to say, numbers have been falsified'

b. *Comme c'est le rapport des experts qui nous conduit à *le* dire, les chiffres ont été truqués.
'As it is the report of experts that leads us to it-say, numbers have been falsified'

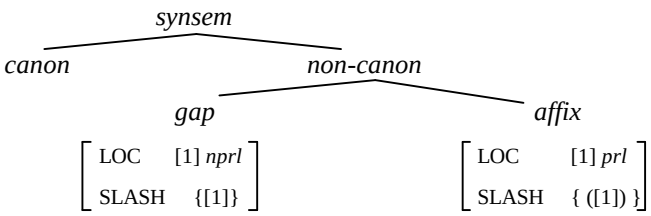
These three standard conditions on extraction contexts in French are uniformly respected by RCCs with both object realizations. Given these results, we conclude that pronominal and gap objects have the same syntactic behavior regarding the RCC unbounded dependency.

Now let us examine hypothesis (1) of RCCs having an object unbounded dependency.

In that hypothesis, we suppose that the gap object is an NP of type *gap-synsem*. That is to say, it has a non-empty SLASH value. By virtue of the Amalgamation Principle (Sag 1997), it follows that the gap object SLASH value would be amalgamated into the slash value of the reportive verb, which triggers the unbounded dependency of the clause.

Since we observed (in examples 5 through 9) that both realizations of the object cause the same general behavior of the construction, it would be consistent to consider the pronominal affix object as a resumptive pronoun that also triggers the dependency. And indeed, it has been already proposed by AGS98 that pronominal affixes—a subtype of *non-canonical* elements—may have a non-empty SLASH value:

(10) typed hierarchy of *synsem*-objects (Abeillé et al. 1998)



Actually, this solution applied to RCCs raises some problems. There exists a general constraint on French relative clauses stipulating that SLASH information must not be passed up beyond the mother node of the adjunct clause. Since RCCs are extractions from adjunct clauses, they must satisfy this constraint too. However, the adverb *comme* cannot be a proper filler for the slashed object. There is a crucial categorial mismatch between the non-canonical NP object and the adverb. Moreover, the semantics of the two are not coreferential. Besides, we exclude the hypothesis of *comme* being a marker, because markers are only of morphological contribution and do not carry any semantic content. Considering the strong semantics of the adverb throughout the different constructions it introduces, it cannot be considered as a marker.

Hence, one problem facing the object extraction hypothesis is the lack of a filler. But, as we will see, the major argument to rule out hypothesis (1) comes from the sensitivity of RCCs to Island constraints.

2.3 Gap vs. resumptive pronoun alternation in French unbounded dependencies

An alternation between gaps and resumptive pronouns can be found in other extraction contexts in French. But, contrary to what happens in RCCs, this alternation leads to significantly different behavior with respect to Island constraints.

We can observe this difference in behavior in topicalization / dislocation constructions, and *dont*-gap relative / *dont*-pronoun relative clauses (on the basis of Godard 1988, and AGS98's analyses). Topicalizations (like in 11a) have a gap argument and they obey Island constraints (11b). On the contrary, dislocations that have a resumptive pronoun (12a) do not obey Island Constraints (12b):

(11) a. Le chocolat, j'aime_.
'the chocolate, I like'

b. *Le chocolat, je sais qui aime_.
'the chocolate, I know who likes'

(12) a. Ce livre, Pierre *l'*a écrit
'this book, Peter it-wrote'

b. Ce livre, je me demande qui *l'*a écrit.
'this book, I to-me-wonder who it-wrote'

As for *dont*-gap relative clauses, they show an NP[*de*] gap argument (ex.13a). They are sensitive to Island constraints (extraction from an embedded relative clause is prohibited, see ex.13b):

(13) a. le livre dont j'ai parlé_
'the book DONT (about-which) I talked'

b. *un auteur dont je connais le critique qui a parlé_
'an author DONT I know the critic who talked'

According to AGS98, *dont*-pronoun relative clauses (DPR) accept a limited set of matrix verbs, mostly propositional attitude predicates. They take a sentential complement which contains a resumptive pronoun (ex.12a). The path between *dont* and this verb is sensitive to Island constraints (14c), whereas the clause containing the resumptive pronoun isn't (14b) (examples 14b-c correspons examples 80 in AGS98:32):

(14) a. Paul_i dont je crois qu'il_i est intelligent
'Paul_i DONT I think that he_i is intelligent'

b. un homme [dont tout le monde savait que le Président n'était pas homme à penser à lui]...
'a man_i DONT everyone knew that the president wasn't one to think of him_i'

c. *un homme [dont il n'y a personne qui sait si le Président va penser à lui]…
'a man_i DONT there's no one who knows if the president is going to think of him_i'

Considering the data (summarized in table 1), we can see that resumptive pronouns are never sensitive to Island constraints[5]. Since RCCs are uniformly sensitive to Island constraints, whether the object is a pronominal affix or a gap, we must conclude that the pronominal affix in RCCs cannot be a resumptive pronoun. Hence, the pronominal affix in RCCs does not trigger any syntactic unbounded dependency, contrary to what is proposed by AGS98 for the pronoun in DPRs.

Table 1 – *Properties of syntactic dependencies in extraction contexts in French*

	Dislocation	Topicalization	*Dont*-gap relative	DPR	RCC
Gap	-	+	+	-	+
Pronominal affix	+	-	-	+	+
Island constraints sensitivity	-	+	+	-	+

In §2.2, we established that given the uniform behavior of RCCs with respect to extraction properties, gap and pronominal objects have the same syntactic status. A further argument is that RCCs have the same reported speech semantics in both cases. Therefore, if the pronominal affix object in RCCs is not involved in a SLASH dependency, then neither is the gap object. It follows that gap objects of reportive verbs must not be of type *gap-synsem*. There is no object extraction in RCCs.

We believe the reported speech semantics of the construction requires there to be an anaphoric relation between the reported string of speech in the main clause and the object argument of the reportive verb. Similarly, the non-canonicity of the object argument as a syntactic and lexical property of reportive verbs is a requirement of the construction. The object argument is anaphoric, but has nothing to do with the extraction dependency.

We propose that RCCs select for a particular class of verbs, a subset of reported speech verbs. Reportive verbs that have a gap object belong to the class of direct speech verbs, as we will see in §3.2.1. We propose that the argument corresponding to the first complement is a particular type of non-canonical synsem.

There exist other cases of absolute complements in French. They are always highly lexically constrained. The predicate OUVRIR (*to open*), for instance, may take an absolute accusative complement, if there is an appropriate antecedent in the previous linguistic context (like in 15a) or a proper source of reference in the deictic context (ex. 15b):

(15) a. Paul a fermé la fenêtre. Plus tard dans la soirée, il a ouvert.

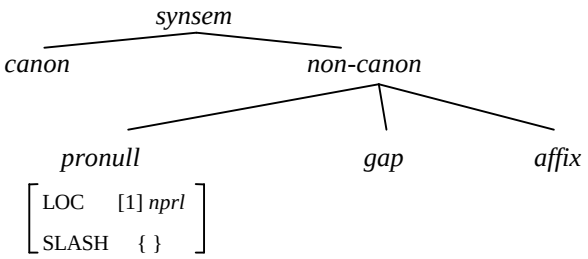
[5] Similarly, there is a correlation between extraction contexts and ETI. ETI only occurs with gap-extractions, never with resumptive extractions.

'Paul closed the window. Later in the evening, he opened.'

b. J'ai tapé trois fois, puis Paul a ouvert.
'I knock three time, then Paul opened.'

This kind of absolute complement is anaphoric. So, we propose to call it *pronull*. Unlike *gap-synsem* and *affix-synsem*, *pronull-synsem* objects always have an empty slash value. Their local value is *non-pronominal* (*nprl*); the typed hierarchy of *non-canonical* objects is as follows:

(16)



As for reportive verbs with pronominal affix objects—a subset of indirect speech verbs—they are cliticized-verbs. Following Miller and Sag's (1997) analysis, these verbs must have an pronominal affix argument on their ARGUMENT-STRUCTURE list. In the case of reportive verbs, the pronominal affix is accusative or dative (it will be later specified as having a *propositional* index).

In order to satisfy the appropriate requirements of reportive verbs, we need a way to ensure that information is correctly passed up along the tree, however far away the mother node of the adjunct daughter may be. Since RCCs are extraction contexts, we will use a SLASH feature. Details of the HPSG treatment we propose are given in §4.

2.4 An adverbial extraction

Long-distance dependencies, ETI, and Island sensitivity prove that RCCs are extraction contexts. Their properties may be fully explained if the extracted element is not the object argument but a gap adverb. *Comme* is thus the filler of an adverbial unbounded dependency. As a dependent of the reportive verb, the slash value of the gap adverb is inherited by the verb, and passed up along the tree until it is 'bound off' by the filler *comme*. Since *comme* and the main clause are not coreferential, we assume that a proper analysis of RCCs is to consider them as a kind of free relative clause: a case of head-adjunct phrase with no coindexation between the modified head and the filler. As a subtype of relative clauses, their internal properties are similar to other gap-relative clauses[6] (like *dont*-gap relative clauses, for instance, they admit ETI and are Island sensitive).

We claim that *comme* is a wh- adverbial word (like *quand—when—*for example, which is a wh-temporal adverb). It belongs to the French wh- paradigm *comme*/*comment*, which is used to express the degree of a property or the manner of a predicate. *Comme* is used to form exclamation clauses (ex.17a), and *comment* forms interrogative clauses (ex.17b):

[6] More precisely, they are a subtype of *gen-qu-rel-cl* (*general-qu-relative* clauses) according to the AGS98 typology.

(17) a. Comme tu es belle aujourd'hui!
'How you are pretty today'
How pretty are you today!

b. Comment prépares-tu les gâteaux?
'How make-you the cookies?'
How do you make cookies?

But, interestingly, we observe that non-exclamative *comme* does not introduce standard gap-relatives. It can never be coindexed with an antecedent in the matrix clause (ex.18):

(18) *J'aime la façon$_i$ comme$_i$ tu parles_
'I like the way how you talk'

Non-exclamative *comme* can only introduce free relatives, and bound gap-adverb dependencies. The semantics of *comme* depends on the semantics of the slashed adverb. In RCCs we believe that it is an adverb marking a reported speech (similar to *ainsi—so—*like in *ainsi l'avait prévu Le Monde*).

Relative clauses are taken to be head-adjunct clauses. Standard relative clauses are only N (or NP) modifiers. We claim that RCCs, as a type of free relative clauses, are also head-adjunct phrases. But unlike standard relative clause, the head they select through the attribute MODIFIED may be different from an N.

In addition, we will show that the paradigm of *comme* adjuncts is an instance of a particular class of French adjuncts which have two types of semantic contribution. They can either be head modifiers, and have a referential contribution to the content of the head; or they can have parenthetical semantics, which means they do not contribute to the referential content of the head. RCCs are analyzed as parenthetical adjuncts introduced by *comme*. We discuss their semantics below.

3 Semantics properties of RCCs

3.1 Parenthetical semantics

As discussed in Schlyter (1997), Espinal (1991), Marandin (1998), and Bonami (1999), some manner adjuncts show a double distribution. They may be predicate modifiers (ex.19b) or be parenthetical adjuncts (ex.19a). In the latter case, they take scope over the whole proposition but do not contribute to the referential content. The most representative adjunct of this class is the adverb *habilement* (*cleverly*):

(19) a. Habilement, Paul détourna la conversation.
'Cleverly, Paul changed the conversation'

b. Paul détourna habilement la conversation.
'Paul changed cleverly the conversation'
Paul cleverly changed the conversation.

We can gloss (19a) by 'the fact Paul changed the conversation was clever' or by 'It was clever from Paul to change the conversation'. Whereas, the gloss in (19b) is 'the process of changing the conversation itself was clever'. A peculiarity of this kind of adjunct is that they change their semantics whenever they change scope. When they modify of the content of a predicate, they are manner adjuncts. But, when they take scope over a proposition, they become parenthetical

adjuncts and take on other meanings. In the case of (19a), for instance, *habilement* is construed as a factive adverb (Schlyter 1977).

We claim that *comme* adjuncts belong to this same class of adjuncts. They can modify predicates, as we have already seen in example (4) when the object argument of the verb is fully realized (and is a repetition of the object argument of the main verb). In this case, they do have the semantics of a manner adjunct. But, they also have parenthetical uses. RCCs are an instance of parenthetical use, with a reported speech meaning.

In addition to exhibiting a parenthetical linearisation[7] (they can be realized in various places), RCCs, as parenthetical adjuncts, do not contribute to the referential content of the head. That means, according to Bouma *et al.* (1999)—henceforth BMS—that they must not appear on its DEPS list. A standard test to show this semantic property is to suppress the parenthetical adjunct. Suppression does not alter the truth conditions of the sentence but only its felicity conditions[8]. Compare (20a) and (20b), adjunction of an RCC do not change the crash event:

(20) a. La bourse s'est effondrée, comme l'avait prévu le NY Times.

b. La bourse s'est effondrée.

Several other tests allow us to reveal the distinctive behavior of parenthetical adjuncts. For instance, whereas parenthetical adjuncts are outsidethe scope of negation (ex.21a), and outside the scope of interrogative operators (ex.21b), manner modifiers are in the scope of negation and interrogative operators (ex. 22):

(21) a. La bourse ne s'est pas effondrée, comme l'avait prévu le NY Times.
The Stock Market didn't crash, as the NY Times predicted.

b. La bourse s'est-elle effondrée, comme l'avait prévu le NY Times?
Did the Stock Market crash, as the NY Times predicted?

(22) a. La bourse ne s'est pas effondrée comme une rangée de dominos
The Stock Market didn't crashed like a row of dominoes.

b. La bourse s'est-elle effondrée comme une rangée de dominos ?
Did the Stock Market crash like a row of dominoes?

The main clause and the RCC are merely semantically appended. Let us use p & q to account for the logical representation of the sentences, where p stands for the main clause and q stands for the RCC in (21) or for the adjunct in (22). (22a) amounts to $\neg(p \,\&\, q)$ which is equivalent to $(\neg p \vee \neg q)$: either the Stock Market didn't crash or it did crash, but not like a row of dominoes. In contrast, as a negative statement, (21a) does not amount to $\neg\ (p \,\&\, q)$ which would mean *either* the Stock Market didn't crash *or* it actually did crash but the New York Times didn't predict it. (21a) always entails that the Stock Market didn't crash, i.e. it amounts to $(\neg p \,\&\, q)$. Similarly, one may utter (22b) while being aware that the Stock Market did crash and just questioning about the way it crashed (the adjunct q). That is not the case in (21b) where the question can only focus on the crash.

[7] See Marandin (1999) for an account of parenthetical linearisation.
[8] More precisely, we posit that RCC contributes a presupposition, see §3.2.

3.2 Reported speech semantics

RCCs impose lexical restrictions on the verbs they admit. The list of verbs includes some predicates of propositional attitude (PENSER – *to think* –, CROIRE – *to believe*), predicates of speech act or communication (DIRE – *to say* –, ÉCRIRE – *to write* –, etc.), and predicates of argumentation (AFFIRMER – *to claim* –, SOUTENIR – *to support* –, CONCLURE – *to conclude* –, etc.). All of them can take a sentential complement. They also share the same semantic argument structure specifications : they realize an AGENT argument as a subject and a THEME argument as an object.

As for the semantics of the construction, we believe that the RCC is an adjunct of discourse reporting. We roughly sketch the speech situation in schema 1. The main speaker (A), using a reportive clause, makes a comment about her own speech. The RCC introduces an embedded discourse context. What is reported is a speech act performed by another speaker (B)—or by the main speaker—but the embedded discourse context is obligatorily different from the main context.

Given P and P' two propositions where P and P' are semantically equivalent,
Speaker A ASSERTS P & Speaker A ASSERTS that speaker B ASSERTED P'

Schema 1 - *Discursive schema for RCC*

Given schema 1, in sentence (16a) P is 'the Stock Market crashed' and speaker B is 'The New York Times'. We use ASSERT as an abstract and very general operator of assertion. It subsumes all the reportive predicates we mentioned before.

An RCC seems to behave as a presuppositional construction. The proposition corresponding to the embedded discourse context (speaker B ASSERTED P) is always true, whatever the truth value of the whole sentence is. That can be evidenced by the test of negative answering:

(23) – La bourse s'est-elle effondrée, comme l'avait prévu le NY Times ?' – Non.
– Did the Stock Market crash, as the NY Times predicted? – No (it didn't).

Usually presuppositions are not supposed to be contested. That is why the negative answer in (23) cannot mean 'No, the NYT didn't predicted the crash'.

In addition, the reportive clause cannot be negative:

(24) *La situation économique s'améliore, comme ne l'a pas indiqué le président.
'The economic situation gets better, as didn't indicate the president'

Considering schema 1, further observations lead us to conclude that RCCs actually have two possible uses: (i) speaker (A) makes reference to a string of speech produced by speaker (B)—in which case A merely quotes B, and the identity of form of P and P' is guaranteed. Or (ii), speaker (A) makes reference to the content of the speech of speaker (B).

This distinction in semantics is supported by empirical data. We show that there are two distinct subtypes of RCCs. We call the first one the *metalinguistic*-reportive clause, and the second one the *propositional*-reportive clause. This distinction provides a semantic explanation of the gap/affix alternation of the object argument of the reportive verb.

3.2.1. **metalinguistic-*reportive clauses (meta-RCC)***

Verbs appearing in metalinguistic RCCs do not lexically realize their objects. The set of predicates is quite limited, it contains DIRE and ÉCRIRE. It is a subset of the set of direct discourse verbs. And indeed, the construction shows direct discourse properties.

The semantic type of the object of metalinguistic reportive verbs is a linguistic sign or a string of linguistic signs which signifiers are guaranteed. That is to say, given schema 1, that P and P' have exactly the same form. And, following Jakobson's (1960) definition of metalanguage, we may say that the content of the object argument has a *metalinguistic* index.

A metalinguistic RCC is a quotative construction. The antecedent of the gap object argument of the reportive verb is understood as a quotation. One consequence of its direct discourse properties is that the antecedent form, or length, is not constrained. It can span the whole main clause (ex.25a), or it can be as short as a lexical word (ex.25b):

(25) a. "Tu vas prendre froid", comme a dit Pierre hier.
You going-to take cold, as said Peter yesterday
You'll get a cold, as Peter said Yesterday.

b. La situation devient vraiment "critique", comme a dit Bob.
The situation becomes really "critical", as Bob said.

A second consequence, as noted by Banfield (1973) is that the quoted antecedent can be in a different language (ex.26):

(26) La situation mérite d'être traitée de façon "politically correct", comme disent les Américains.
(equivalent to : The situation deserves to be treated in a "politiquement correcte" way, as French people say.)

Whatever the object antecedent is, it is always the element in the scope of the metalinguistic-RCC adjunct.

There is a strong similarity between Direct speech arguments and quoted arguments. In both cases, the string uttered must be construed with its exact form. As we know, quotation and direct speech reporting represent a frequent linguistic activity. But representing the selection involved in these constructions (via a specifier, predicate, or adjunct) is theoretically problematic. Linguistic elements are usually described in terms of phonology, syntax, and referential semantics, but what we need in the metalinguistic case is a mechanism for selecting linguistic signs themselves.

Selection cannot rely on any particular categorial restrictions, since any sign may be quoted or be a direct speech report. Similarly, we cannot rely on semantic selection of a particular entity in the world. Metalinguistic elements are somehow construed on a double level. They denote signs as objects of the world, but at the same time, inside the quotation or the report clause, they can construct and contribute referential semantics. Given that selection mechanisms in HPSG handle objects whose maximal sort is *synsem*, we believe that every *synsem* object must contain a metalinguistic value in addition to the rest of its description. This value must be unique in order to individuate the instantiation of a *synsem* object. Therefore, we propose the following description for every object of type *synsem*:

(27)

$$\textit{synsem}\begin{bmatrix}\text{META-INDEX} & \textit{meta-index}\\ \text{LOCAL} & \textit{local}\\ \text{NON-LOCAL} & \textit{non-local}\end{bmatrix}$$

Every *synsem* object must have a non-empty value for the attribute META-INDEX. The value of feature META-INDEX must be an object of sort *meta-index* (*meta-ind*).

Specifiers, predicates or adjuncts that select a metalinguistic argument will specify the META-IND value of the SYNSEM of this argument. If they don't select a metalinguistic argument, then there is no need to constraint the META-IND value of the selected argument. This META-IND feature allows the grammar to cover all cases of metalinguistic argument selection.

In the case of meta-RCCs, all verbs entering the construction will specify the *meta-ind* value of their *pronull* argument.

Let us turn now to propositional-RCCs.

3.2.2. propositional-*reportive clauses (prop-RCCs)*

Verbs of propositional RCCs realize a pronominal affix object which is construed as a sentential argument. The set of predicates entering the construction is quite large. It represents a subset of the set of indirect discourse verbs. The object pronoun is unmarked for agreement; its form is *le* (accusative case) or *y* (dative case)—determined lexically based on the form of the sentential complement subcategorized for by the verb when it realizes it.

In the *propositional*-reportive clause, the object antecedent corresponds to the whole main clause. Thus, the pronominal affix object must have a *propositional* index.

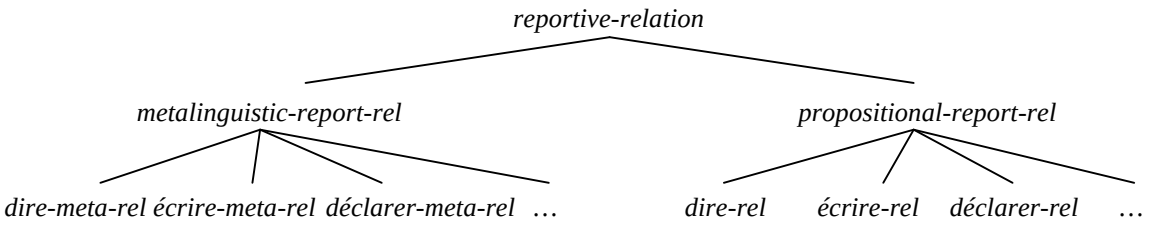
4 Analysis

4.1 Reportive verbs

The semantic relation of reportive verbs needs to be specified according to the type of object argument they select for. We propose a general typed hierarchy of reportive semantic relations which separates direct speech/quotation from indirect speech relations, as sketched in (28)[9]:

[9] Actually only DIRE and ÉCRIRE appear in meta-RCCs, therefore we would need a further subtype in order to avoid other direct speech verbs.

(28)



Predicates with a *metalinguistic-report* relation specify a value for the META-IND of their object argument, while predicates with a *propositional-report* relation require that the CONTENT | INDEX value of their object argument to be of type *propositional*.

This distinction is crucial for reportive verbs in RCCs. It allows us to avoid ambiguity between verbal lexemes. For instance, lexeme DIRE in RCCs will have two different entries according to the type of its reportive relation. Consider the instantiation of the words *dit* and *le-dit* when they appear respectively at the bottom of a meta-RCC (29) and a propositional-RCC extraction context (30):

(29)

$$\begin{bmatrix} \textit{dit} \\ \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{DEPS-LIST} & \langle\ \text{NP}\textit{nom}\ [\ \text{IND}\ [0]\ ],\ [1]\textit{nullpro-ss}\begin{bmatrix} \text{CAT} \mid \text{CASE} & \textit{acc} \\ \text{META-IND} & [2] \end{bmatrix},\ \text{ADV}\textit{gap-ss}\ [\text{LOC}[3]],\ \ldots\ \rangle \end{bmatrix} \\ \text{LOC} \mid \text{CONT} \begin{bmatrix} \textit{meta-dire-rel} \\ \text{AGENT} & [0] \\ \text{THEME} & [2] \end{bmatrix} \\ \text{NLOC} \mid \text{SLASH}\ \{[3]\} \end{bmatrix}$$

(30)

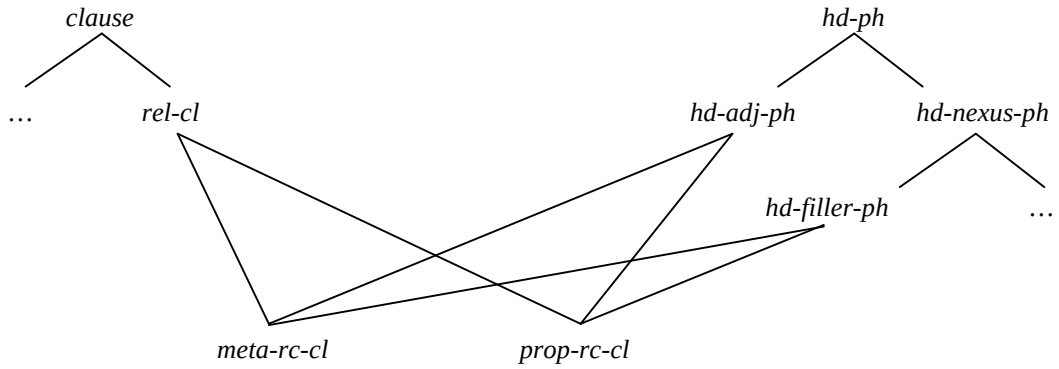
$$\begin{bmatrix} \textit{le-dit} \\ \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD} & \textit{cl-verb} \\ \text{DEPS-LIST} & \langle\ \text{NP}\textit{nom}\ [\ \text{IND}\ [0]\ ]\ ,\ [1]\ \textit{praff-ss}\begin{bmatrix} \text{CAT} \mid \text{CASE} & \textit{acc} \\ \text{CONT} \mid \text{IND} & [2] \end{bmatrix},\ \text{ADV}\textit{gap-ss}\ [\text{LOC}[3]],\ \ldots\ \rangle \end{bmatrix} \\ \text{LOC} \mid \text{CONT} \begin{bmatrix} \textit{prop-dire-rel} \\ \text{AGENT} & [0] \\ \text{THEME} & [2] \end{bmatrix} \\ \text{NLOC} \mid \text{SLASH}\ \{[3]\} \end{bmatrix}$$

The word instantiated in (29) may appear for example in sentences like *comme dit le NY Times ce matin* (lit: 'as says the NY Times this morning'), and the one in (30) may appear in sentences like *comme le dit le NY Times ce matin* (lit: 'as it-says the NY Times this morning').

4.2 RCCs in the cross-classified typed hierarchy

Following Sag (1997) and AGS98 for French, phrases are organized along two dimensions: clausality and headedness. Each clause is defined with respect to its clausal type(s) and its phrasal type(s). RCCs are a subtype of *relative-clause* and is an instance of *head-adjunct* phrase and *head-filler* phrase. The classification of meta-RCC and prop-RCC is as follows:

(31)



The reportive-*comme* construction is a head-adjunct phrase, where the *comme*-clause is the adjunct-daughter (i.e. the nonhead-daughter—abbreviated NHD-DTR). The SYNSEM value of the head-daughter phrase (HD-DTR) is selected by the adjunct phrase through the attribute MODIFIED (MOD). But, unlike in relative clauses, the category of the head-daughter is underspecified. According to AGS98[10] head-adjunct phrases must share their CONTENT specifications with those of the adjunct daughter. Therefore, RCCs inherit the following constraints from head-adjunct phrases:

(32)

$$\textit{hd-adj-ph} \Rightarrow \begin{bmatrix} \text{CONT} & [1] \\ \text{HD-DTR} \mid \text{SYNSEM} & [3] \\ \text{NHD-DTRS} & \left\langle \begin{bmatrix} \text{HEAD} \; [\text{MOD} & [3]\,] \\ \text{CONT} & [1] \end{bmatrix} \right\rangle \end{bmatrix}$$

The adjunct-phrase in a RCC is a case of filler-gap extraction, hence it must satisfy the following specifications from head-filler phrase:

[10] This constraint corresponds to the Semantics Principle in Pollard and Sag (1994).

(33)

hd-filler-ph ⇒

$$\begin{bmatrix} \text{SLASH} & [1] - \langle [2] \rangle \\ \text{HD-DTR} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SLASH} & [1] \end{bmatrix} \\ \text{NHD-DTRS} & \langle [\ \text{LOC}\ [2]\ ] \rangle \end{bmatrix}$$

4.3 The gap-adverb LOCAL value

The gap-adverb of the reportive clause plays a key role in our analysis. The dependents-list (DEPS-LIST) of reportive verbs that appear in RCCs must record the presence of an adverb of type *gap-synsem*. Following the Slash Amalgamation Constraint and the Slash Inheritance Principle (Sag 1997, BMS) the verb will store the SLASH value of this dependent into its own SLASH set value, and will pass it up along the tree until it is identified and bound off by the local value of the filler-daughter; i. e. *comme*.

The designation of the appropriate set of verbs relies on a constructional constraint. Still, it raises some difficulties. As we observed, RCCs may be long-distance dependencies, and the reportive verb is always in the lowest clause. In order to ensure that the lowest verb is a member of the appropriate set of verbs, without violating locality conditions of wellformedness, we would need a local way to select it.

Now, reportive verbs in RCCs are selected via the MOD value of the gap-adverb. Since the LOCAL value of the gap-adverb is the one propagated by the SLASH information throughout the path and is token identical to the LOCAL value of the filler-daughter, constructional specifications are simply encoded in the local description of the filler-daughter.

At this point, we can define the two RCC constructions:

(34)

meta-rc-cl ⇒

$$\begin{bmatrix} \text{NHD-DTR} & \begin{bmatrix} \textit{comme} & \\ \text{MOD} & \begin{bmatrix} \text{CONT} & \textit{meta-report-rel} \\ \text{DEPS} & \langle \ldots, \begin{bmatrix} \textit{pronull} \\ \text{META-IND}\ [1] \end{bmatrix}, \ldots \rangle \end{bmatrix} \end{bmatrix} \\ \text{MOD} & [\ \text{META-IND}\ [1]] \end{bmatrix}$$

(35)

prop-rc-cl ⇒

$$\begin{bmatrix} \text{NHD-DTR} & \begin{bmatrix} \textit{comme} & \\ \text{MOD} & \begin{bmatrix} \text{CONT} & \textit{prop-report-rel} \\ \text{DEPS} & \langle \ldots, \begin{bmatrix} \textit{praff} \\ \text{CONT} \mid \text{IND}\ [1] \end{bmatrix}, \ldots \rangle \end{bmatrix} \end{bmatrix} \\ \text{MOD} & \begin{bmatrix} \text{S} \\ \text{CONT}\ [1]\ \textit{proposition} \end{bmatrix} \end{bmatrix}$$

The obligatory anaphoric relation is captured by token identity between (1) the CONT | INDEX values of the head-daughter of the head-adjunct phrase and the object argument of the reportive verb (as in 35); or between (2) the META-INDEX values of the head-adjunct head-daughter the object argument of the reportive verb (as in 34).

Constraints on extraction propagation are directly inherited from the *head-filler-ph* type. Given the local description of filler *comme*, the appropriateness of the gap-adverb and the verb can be verified.

4.4 A wrinkle: parenthetical adjunction

Given constraints on head-adjunct phrases, the CONTENT value of the RCC head-daughter should be shared with the one of the head-adjunct mother node. Actually, this is inconsistent with the non-referential contribution of the parenthetical. Because if a head-adjunct phrase containing an RCC is itself selected by any predicates, then the semantic content of the RCC would be visible in the referential content of the predicate.

The solution we adopt here, in order to block the referential contribution of RCCs, is to divide head-adjunct phrases into two subtypes : *head-modifier-adjunct phrase* and *head-parenthetical-adjunct phrase*. It is only in the former subtype that the adjunct-daughter is a semantic head for the head-adjunct phrase. In the *head-parenthetical* subtype, the syntactic head daughter is also the semantic head daughter of the phrase.

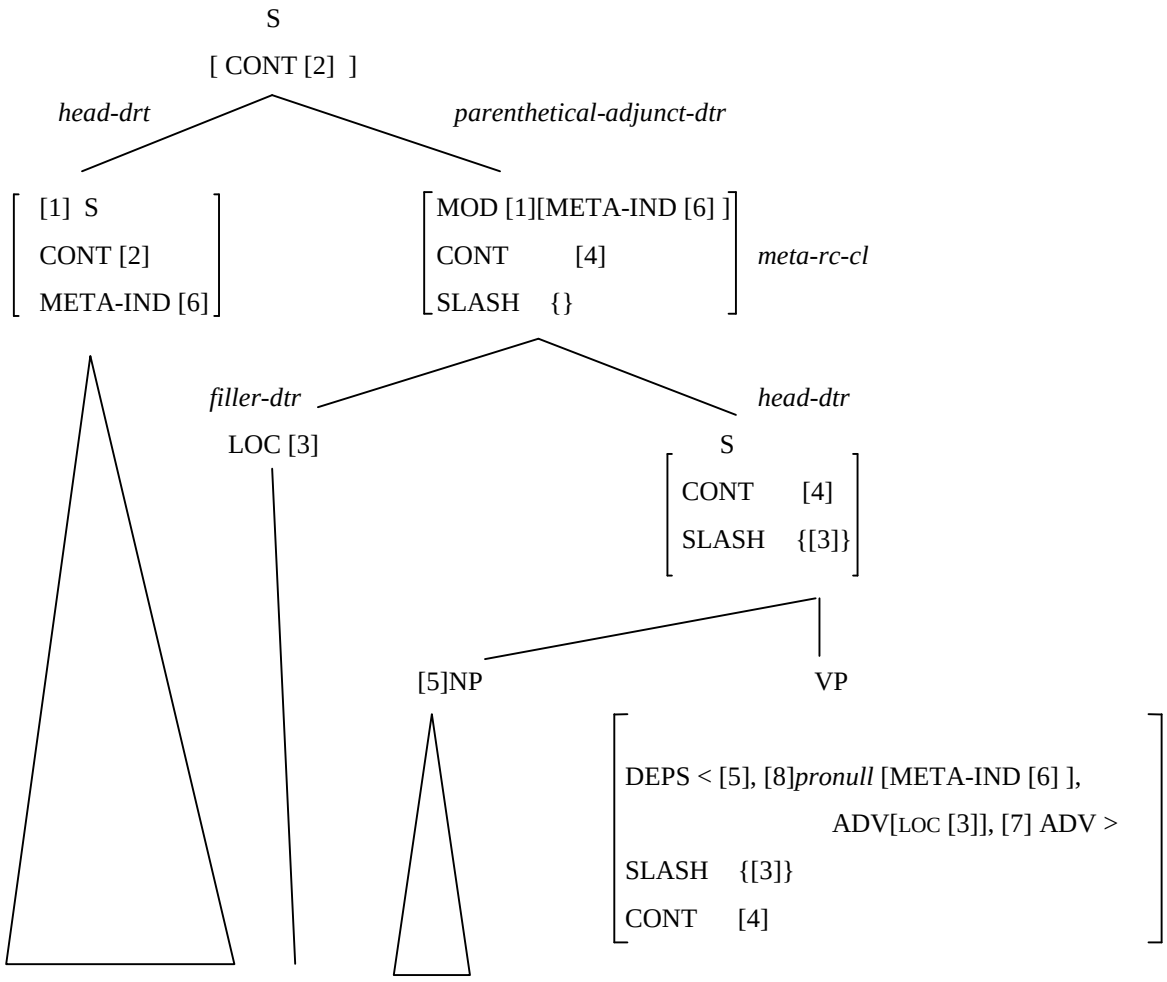
4.5 Representation

We put all our proposals together by giving illustrations of the two distinct reportive-*comme* constructions. The representations corresponding to sentences (36) and (37) are given respectively in (38) and (39).

(36) “La situation est critique”, comme le président a dit hier.
“The situation is critical”, as the President said yesterday.

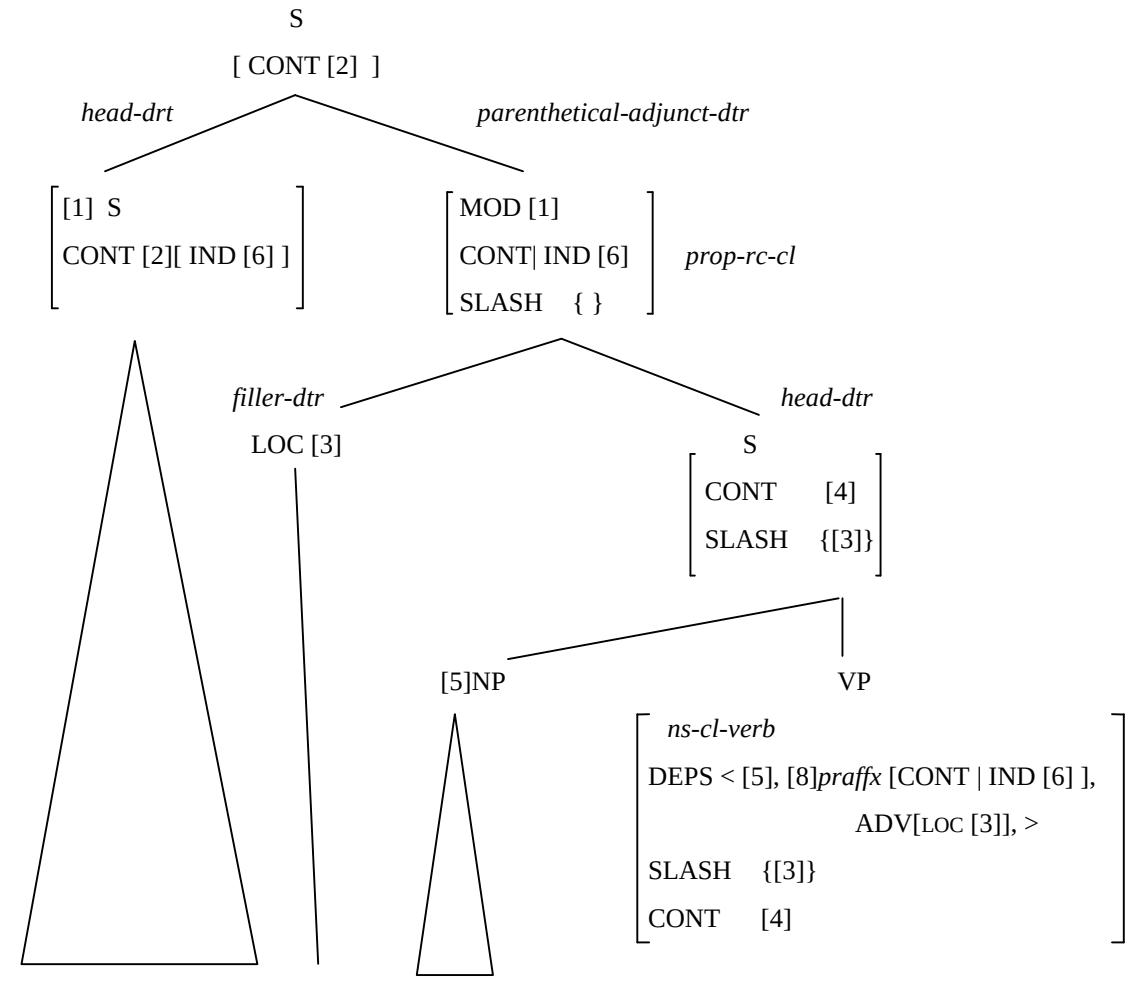
(37) La bourse s’est effondrée, comme Le Monde le prévoyait.
The Stock Market crashed, as Le Monde predicted it.

(38)



"La situation est critique", comme le président a dit hier

(39)



5 Conclusion

In this paper, we presented a new set of data to be integrated in a surface-based grammar for French. We showed that a particular use of adjunct clauses introduced by the adverb *comme*—that we called *reportive-comme* clauses (RCC)—exhibits a large amount of distinctive properties such that it is justified to consider it as a construction, on its own. Following Sag (1997) and Abeillé *et al.* (1998) which rely on the central notion of 'construction' for the classification of grammatical phrases, our analysis of RCCs permits to embody their linguistic complexity in terms of constructional constraints. RCCs are part of a cross-classified typed hierarchy. As adverbial extraction contexts they are a subtype of the relative clauses clausal type and they inherit a large part of their syntactic and semantic constraints from the head-adjunct and head-filler phrasal type. Considering long-distance dependencies RCCs may admit, the adverbial SLASH information plays a key role for the appropriate selection of reportive verbs and for the representation of the extraction dependency. Our treatment allows also to express that semantic

specifications of RCCs are realized along different dimensions. We drew a general distinction between head-modifier adjuncts and parenthetical adjuncts in order to account for the fact that parentheticals do not contribute the referential content of the head-phrase they select for. We posited two subtypes of RCCs determined by a Direct speech (and quotative) vs. Indirect speech distribution of properties. The sets of defining constraints for these two subtypes allow us to characterize the restricted classes of verbs possible in the different RCCs, the syntactic realization (gap or pronominal affix) of their object argument and its anaphoric semantics. We also considered a particular case of non-canonical argument realization which does not involve any extraction, and proposed the enrichment of the *non-canonical-synsem* typed hierarchy with a *nullpro* subtype. This can also provide an adequate treatment for non-realized arguments of other constructions, like in quotative inversion.

The main point of our analysis is to propose a formalized account for direct speech or quoted argument selection, introducing the META-INDEX feature on every *synsem* object and a typed hierarchy for semantic relations of reported speech predicates. Direct reported speech phenomena are known to be a puzzle of the syntax-semantic interface that formal syntactic theories rarely account for. The constraint-based representations we propose here capture properly some aspects of this interface complexity.

Our analysis emphasizes on one of the numerous uses of *comme*-adjuncts. It is based on larger work (Desmets 2001) that elaborates a unified treatment for the major *comme*-constructions of French grammar.

References

ABEILLÉ, A., D. Godard & I. A. Sag. 1998. *French Relative Clauses*. Ms. University Paris 7.

BANFIELD, A.1973. Grammar of Quotation, Free Indirect style, and the Implications for a Theory of Narrative, *Foundations of Language* 10: 1-39

BONAMI, O., D. Godard & J.-M. Marandin, 1999. "Constituency and word order in French subject inversion", *In* Gosse Bouma, Erhard Hinrichs, Geert-Jan M. Kruijff et Richard T. Oehrle (eds.), *Constraints and ressources in natural language syntax and semantics*. Stanford : CSLI Publications.

BONAMI, O., D. Godard & J.-M. Marandin, 1998. French subject inversion in extraction contexts, *In* Gosse Bouma, Geert-Jan M. Kruijff, et Richard T. Oehrle, (eds.), *Proceedings of FHCG'98*, pp. 101-112. Saarbrücken.

BONAMI, Olivier, 1999. *Les constructions du verbe : le cas des groupes prépositionnels argumentaux*, PhD - Paris 7 University.

BOUMA, G., R. MALOUF & I. A. SAG. 1999. *Satisfying Constraints on Extraction and Adjunction* (Kluwer Academic Publishers - to appear).

DESMETS, M. 1998. Identification de deux constructions en *comme* : causalité et comparaison, *LINX* 39-2: 89-118

DESMETS, M. 2001. *Les typages de la phrase en HPSG : le cas des phrases en "comme*". Doctorat thesis, Univ. Paris X. (to appear).

ESPINAL, T. 1991. The representation of disjunt constituents, *language* 67-4: 726-762

GODARD, D. 1988. *La syntaxe des relatives en français*. Editions du CNRS

HALL PARTEE, Barbara, 1973. The syntax and semantics of quotation, *In* Stephen R. Anderson and Paul Kiparsky (eds), *a Festschrift for Morris Halle*. pp. 410-418. Holt, Rinhart and Winston Inc.

HUKARI, T. E. & R. D. Levine.1995. Adjunct extraction, *Journal of Linguistics* 31 :195-226.

JAKOBSON, R. 1960. Linguistics and Poetics, *In Style in Language*, T.A. Sebeok (reed.), Cambridge: MIT. Ed

KAYNE, R.1973. L'inversion du sujet en français dans les propositions interrogatives, *Le français moderne*, pp 10-151.

KAYNE, R. & J.–Y. Pollock.1978. Sytlistic Inversion, Successive Cyclicity and Move NP in French, *Linguistic Inquiry* 9: 595-621.

LAPOINTE, G. Steven, 1991. Empty Categories in Reportive *As* Clauses, *First Meeting of Formal Linguistics Society of the Midwest (FLSM)*, University of Wisconsin-Madison.

MARANDIN, J. – M. 1998. *Grammaire de l'incidence*. Ms. University Paris 7

MILLER P., I. A. Sag, 1997. French clitic movement without clitic or movement, *Natural Language and Linguistic Theory.* 15: 573-639.

POLLARD, C. & I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar,* Chicago: Chicago University Press

REAPE, M.1994. Domain union and word-order variation in German, *In* J. Nerbonne, K. Netter et C. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, pp. 151-197. CSLI : Stanford

ROSS J. R., 1967. *Constraints on variable in syntaxe*, PhD, MIT.

SAG, I.A.1997. English relative clause constructions, *Journal of Linguistics* 33: 431-484

SCHLYTER, S. 1977. *La place des adverbes en -ment en français*, ms. Constance univ.

WEBELHUTH, Gert, 1992. *Principles and Parameters of Syntactic Saturation*. New York : Oxford University Press.

ZWICKY, Arnold M., 1974. *Direct and Indirect Discourse*, WPL 17 : 198 –205.