THE DRAMATIC EXTRACTION CONSTRUCTION IN FRENCH

Anne Abeillé, Danièle Godard, Frédéric Sabio

Abstract: Relying on spoken corpora (Corpaix, CRFP) and on previous studies (Sabio 1995, 2006), we identify a construction common in spoken French, which we analyze as a particular case of extraction:

a. *dix sept ans* il a. (Seventeen years he has) [Corpaix]
b. *deux cigarettes* j'ai fumé. (Two cigarettes I smoked) [on the fly]

The construction can only be a root clause and a declarative clause. Its interpretation is that of a thetic proposition. On the other hand, it is not associated with a unique information structure, since it is compatible with a focus-ground partition, with the extracted constituent as a narrow focus, or with an all focus interpretation. We call this construction "dramatic extraction", and the extracted element a "center" (i.e. a focus or a figure). We formalize our analysis within the HPSG framework.

Keywords: French, syntax, focus fronting, Romance, HPSG

1. Introduction

It had long been assumed that French lacked the possibility of extracting a constituent in a declarative clause, thus contrasting for instance with English, where NP topicalization is well-known. However, the empirical basis taken into account was too narrow: in particular, constructions which are specialized for spoken everyday speech were not taken into account. The collection of spoken data has been rendered possible by the new technologies; although the effort has to be pursued to arrive at a large data base, and more congenial interfaces, some corpora of spoken French are now available (such as Corpaix, Elicop, CRFP). It is clear that extraction of an NP is in fact available in spoken French (see Sabio 1995, 2006); furthermore, while the construction is restricted, regarding the register and conditions of use (spoken everyday speech), it is in no way socially marked, so that it is possible to appeal to speakers' intuitions, to better understand its properties. We thus appeal to corpora as well as intuitions, and (short) examples taken on the fly. In particular, we use two corpora collected in Aix-en-Provence (Corpaix and CRFP).[1]

After showing that the preposed element is extracted (as opposed to dislocated), we turn to the semantic and pragmatic properties that characterize the construction. Because of its dialogical properties, we call it the 'dramatic extraction construction'. Finally, we briefly indicate how the construction can be described in the syntagmatic grammar HPSG.

2. The preposed constituent is extracted

Consider the sentences in (1), which contain a preposed NP.[2]

[1] Corpaix is a spoken corpus of more than 1 M words collected in the 90's in Aix-en-Provence, consisting mostly of interviews (Blanche-Benveniste et al. 2002); CRFP (Corpus de référence du français parlé) of 440 000 words, consisting mostly of interviews, recently collected in major French cities (http://www.up.univ-mrs.fr/delic/crfp/). Pohl (1984) has collected his own data on the fly.

[2] Following common practice, we do not use punctuation marks for the examples taken from spoken corpora, except for the major ones (dot, question or exclamation mark). '//' indicates the presence of a perceptible pause. Throughout the text, we give rough transpositions in English, rather than gloses and translations. We indicate the initial NP with italic characters.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

(1) a. *dix sept ans* il a. (Seventeen years he has) [Corpaix]
b. *deux cigarettes* j'ai fumé. (two cigarettes I smoked) [on the fly]

The relation between the preposed NP and the rest of the sentence has the properties of extraction. First, the preposed NP corresponds to an unrealized NP in the sentence, whose grammatical function it takes on. In (1), the NP corresponds to an object. But the NP can correspond to unrealized categories with different functions: a predicative phrase (2a), a specifier (2b), an oblique NP denoting a localization in (2c)[3]. In all cases, the sentence with the NP *in situ* would be grammatical (3). The NP is often obligatory (4):

(2) a. *Chirurgien* elle était. (Surgeon she was)
b. *Trois heures* il avait de retard, le train! (Three hours it had of delay, the train)
c. *Place de la nation* on a été. (Nation square we went (to))
(3) a. Il a dix-sept ans. (He is seventeen year old)
b. J'ai fumé deux cigarettes. (I smoked two cigarettes)
c. Il avait trois heures de retard, le train! (It had three hours of delay, the train)
d. On a été Place de la Nation. (We went (to) Nation Square)
(4) a. * Elle était. (She was)
b. J'ai fumé. (I have smoked)
c. *Il avait de retard le train! (It had of delay the train)
d. *On a été. (We went (to))

Second, there is a long distance between the preposed NP and the unrealized category. Thus, the dependency can cross a sentential barrier.

(5) a. *Dix-sept ans* je crois qu'il avait à l'époque.
(seventeen years, I think that he was at the time)
b. *Deux cigarettes seulement*, je te dis que j'ai fumé.
(Two cigarettes only, I tell you that I smoked)
c. *Chirurgien*, j'ai lu qu'elle était à l'époque.
(Surgeon, I read that she was at that time)
d. *Trois heures*, je te dis qu'il avait de retard, le train!
(Three hours, I tell you that it had of delay, the train)
e. *Place de la Nation*, on a appris qu'ils étaient allés.
(Nation Square, we learned that they had gone (to))

Third, the path between the NP and the unrealized category obeys island constraints. For instance, we see that the unrealized category cannot belong to a relative clause, a weak island (such as an embedded interrogative clause) or an adjunct clause.

(6) a. **Dix-sept ans*, je ne vois personne [qui ait ici.]
(seventeen years, I see nobody who is around here)
b. ?*Chirurgien*, on se demande [si elle est vraiment.]
(Surgeon, we wonder whether she really is)

[3] We follow the terminology and analyses of the *Grande Grammaire du français* (forthcoming) to be published by Bayard Editions in 2010. For the analysis of measure phrases followed by a *de*-marked N as specifiers + heads, see e.g. Milner (1978), Abeillé et al. (2004). We analyze NPs (not included in a PP) denoting places and times as oblique complements. For a systematic comparison with another construction where the preposed NP is analysed as left dislocated, see Abeillé et al. 2008.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

c. **Trois heures*, je suis parti à la gare [sans savoir qu'il avait de retard, le train]!
(Three hours, I went to the station without knowing that it had of delay, the train)

Although we concentrate here on extracted NPs, it must be noted that the construction is not restricted to NP fillers, and also exists for APs, (certain) PPs, (certain) adverbs and VPs.

(7) a. *A moitié anglaise*$_{AP}$ elle était. (Half British she was) [Arletty, Corpaix]
b *A une sorcière*$_{PP}$ tu ressembles. (To a witch you resemble)
c *Gentiment*$_{AdvP}$ il s'est comporté. (Kindly he behaved)
d *Le laver*$_{VP}$ il faut. (Wash it one must) [Pohl]

Finally, the sentence (following the extracted constituent) can begin with the complementizer *que*, in non standard French (noted with !). It is a common feature for the head sentence in extraction constructions (relative or interrogative clauses) to begin with such a complementizer in non standard French (8c, d).

(8) a. !*Six euros* que ça m'a coûté! (six euros that it cost me)
b. !une fois j'ai mis du vert – *affreux* que c'était
(once I wore green, horrible that it was) [Pohl]
c. !là où que c'est brûlé (there where that it is burnt) [Corpaix]
d. !Qui que tu as vu ? (Whom that you saw?)

3. Semantic and discursive properties

In spite of being syntactically made of two parts, the construction is interpreted as a thetic proposition. Moreover, it is not always associated with the same information structure, being either partitioned into ground and a narrow focus (corresponding to the extracted constituent) or constituting an all focus utterance.

3.1 The content of the sentence is a thetic proposition

As seen in the preceding section, the construction is syntactically partitioned into an extracted phrase (a filler) and a sentence missing an XP. Interestingly, this syntactic partition is not matched with a partitioned proposition, but with a thetic one. A categorical proposition with a salient NP would be partitioned; however, the content of the construction under study does not have the required or typical properties of a categorical proposition[4]. In a categorical proposition, a property is predicated of an entity (to which an argument refers) while a thetic proposition describes a situation. A categorical proposition is about an entity, while a thetic one is about the situation with which the whole sentence is associated.

The salient phrase of a categorical proposition (or sentence Theme, Marandin 2007) can be recognized both by formal and referential properties. First, if we think of the content of the verb as a predicate with arguments places or variables that are filled with contentful arguments as the verb combines with argument phrases; the salient phrase of a categorical proposition corresponds to the highest leftmost term, in a language such as English, German or French. This corresponds to the notion of a "semantic subject" in Jacobs (2001), taken up in a formal setting by Webelhuth (2007). In addition, the sentence Theme has an affinity with

[4] See e.g. Ladusaw (1994), Kim (1998), for arguments to the effect that the categorical vs. thetic distinction is semantic rather than discursive or pragmatic.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

definite NPs, because its referent is preferably discourse familiar (Prince 1981). Finally, it combines preferably with an individual level predicate (see Ladusaw 1994).[5]

The extracted NP in the construction illustrated in (1) has some formal properties of the sentence Theme: being extracted and initial, it is the leftmost highest phrase with which the sentence combines. But it does not necessarily correspond to a semantic argument. Not only can it be an oblique (2c), but it can also be the specifier of an NP (2b), and it can be itself the predicate (2a). Furthermore, the NP in (1) certainly lacks the referential properties of a sentence Theme: indefinite NPs and non referential phrases (predicates) are preferred, idiom chunks are allowed (11c) and the predicates they combine with are not constrained in terms of the individual vs stage level distinction.

We illustrate in (9) the construction with indefinite NPs referring to a particular entity.

(9) a. et là, tu sais ce qui lui est arrivé // *une antenne* ils lui ont jeté sur la tête !
[Corpaix : Nord, 40, 3]
(And then, do you know what happened to him? An antenna they threw on his head)
b. Sp A. Donne-moi un redonex. (Give me a redonex)
Sp B. *Un redonex* tu veux ? [Pohl (1984)] (A redonex you want?)
c. Mon père il va m'acheter un petit mouton // *un petit mouton* il va m'acheter.
[Corpaix: Agenet, Gr3, 6]
(My father he is going to buy me a small sheep // a small sheep, he is going to buy for me)

More typically, the extracted NP refers to different types of objects, measures in particular: the NP measures the duration of an event (10a), or its frequency (10b), or it estimates the age of a person (1a). A predicate is also frequently extracted as in (11a, b), where *faire horreur* is a support verb construction (hence the noun is part of the predicate), and an idiom chunk is possible too (11c).

(10) a. *Onze heures* elle est restée chez les juges ! [*Canard Enchaîné*, 2006]
(eleven hours she remained with the judges)
b. Tu l'as pas vu une seule fois aux informations // *pas une fois* tu l'as vu.
[Corpaix: Nord, 40, 3]
(you didn't see it once on TV, not once did you see it)
c. J'ai commencé à sept ans // sept ans et demi huit ans je sais plus exactement quatre-vingt-s *huit ans* je devais avoir [CFRP: BOR-R00PRI001]
(I started at seven // seven and a half I don't remember exactly eighty-s eight years I must have been)

(11) a. J'ai écrit dans le journal local d'Aire-sur-la-Lys je me rappelle plus maintenant ah *L'Echo de la Lys* // ça s'appelait je crois bien. [Corpaix]
(I wrote in the local newspaper of Aire-sur-la-Lys I don't remember now ah L'écho de la Lys it was called I think)
b. *Horreur*, je lui faisais, docteur. [R. Forlani, *Ma chatte ma folie*, 1992:15]
(Horror she had of me, doctor)
c. *Des clopinettes* il m' a donné (peanuts he gave me)

[5] For the distinction between individual and stage level predicates, see Kratzer (1995).

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

Definite NPs are not impossible, but they are certainly dispreferred. Here are some made-up (but plausible) examples.

(12) a. Tu sais ce qui est arrivé ? *Le candidat du patron*, ils ont refusé!
(You know what happened ? The boss' candidate they did not accept!)
b. Sp. A. Je cherche mes lunettes. (I am looking for my glasses)
Sp. B. *Tes lunettes*, tu cherches? (Your glasses you are looking for?)

Thus, we can conclude that the extracted NP does not function as a sentence Theme in a categorical proposition. Accordingly, we do not have a categorical, but a thetic proposition, which describes a situation as a whole, and does not select a semantically salient entity.

3.2 Discourse properties

3.2.1 Not a segment of a narration

A striking property of the construction is that it is not a segment of a narrative discourse. In particular the NP is not easily taken up by a pronominal. Thus, the sentence in (13) is not followed by an explanatory sentence containing a pronoun referring directly to the entity denoted by the NP (13a), although it can be followed by an explanation concerning the whole situation or speech act (11b). In (13a) the pronoun *ils* refers to the proofs of residence, while the general pronoun *ce* in (13b) refers to the request ((13b) can be paraphrased as 'I am making this request because I need them in order to be registered as a voter'). This is also true when the sentence itself could easily appear as belonging to a narrative sequence : (9a) describes an unexpected event, which follows other events. However, it cannot be followed by a sentence describing the result of the action on the state of the entity (14), although, again, it can be followed by an assessment of the situation.[6]

(13) *Deux justificatifs de domicile*$_i$, il me faut. [Pohl (1984)]
a. # Ils$_i$ sont nécessaires pour l'inscription sur les listes électorales.
b. C'est pour l'inscription sur les listes électorales.
(Two proofs of residence I need. They are necessary for my registration on the voters' lists. / It's for my registration on the voter's lists).
(14) et là, tu sais ce qui lui est arrivé // *une antenne*$_i$ ils lui ont jeté sur la tête!
a. #Elle$_i$ était complètement fichue!
b. C'est incroyable!
(And then, do you know what happened to him? An antenna they threw on his head! It was completely ruined / It is incredible)

This confirms that the sentence is not about the entity denoted by the extracted NP. It also points towards a special illocutionary status of the sentence (see below section 4).

3.2.2 Information Structure properties

In terms of the focus vs. ground distinction, there is no unique information structure associated with the construction in (1), but (at least) two different possibilities. Let us think of a text, discourse or conversation as a series of utterances, each of which answers one of the question raised by the preceding context (see e.g. Büring 1997, Ginzburg 2008, Roberts 2004). The focus is that part of the content which contributes novel information, thus

[6] "#" notes discursive or pragmatic inappropriateness.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

answering a contextual question. The extracted NP can constitute a narrow focus, as in (11), where the NP 'l'Echo de la Lys' answers a question raised by the context; more precisely, in this case, the question under discussion has been partially expressed by the speaker (it is interrupted by the answer).

The construction is not compatible with a narrow focus on another element, as shown by the unfelicity of the following answers:

(15) Sp.A. Quand est-ce qu'ils lui ont jeté une antenne sur la tête?
Sp. B. #*Une antenne* ils lui ont jeté sur la tête hier.
(When did they throw an antenna to his head? An antenna they threw to his head yesterday)

(16) Sp.A. Qu'est-ce que tu as attendu deux heures?
Sp.B. #*Deux heures* j'ai attendu le bus
(What did you wait for 2 hours? Two hours I waited for the bus)

Alternatively, the whole sentence can be all focus. This is particularly clear in (9a), where the sentence answers the typical question for an all focus sentence: 'and then what happened?', which is expressed by the speaker. The sentence given in (10a), constitutes the caption under a satirical drawing (showing a minister being interrogated by the judges); there are thus good chances that the whole drawing with its caption constitutes new information (in the context).

Lastly, we should point out the interesting but unclear case illustrated in (9c) and (10b). Our construction appears as the second sentence of a pair, the first of which already gives the whole content of the second. We can see the second sentence as being a pure repetition, at least informatively, thus an all ground sentence. But we can also think of it as being a confirmation or a reassertion, in which case it would be here like an all focus sentence.

The fact that a given syntactic form is not matched with a unique focus-ground partition has already been pointed out. Prince (1998, 1999), in particular, has been a ground breaker in this domain. While NP preposing in English had been assumed to be linked with a topic status for the filler (hence, the name "topicalization" for the construction), she shows that the focus can be associated with either the filler or the sentence head (the constituent whose content represents the focus is underlined in (17), as it is in Prince 1998, 1999).

(17) a. Let's assume there's a device which can do it – <u>a parser</u> let's call it. What follows? [J. Fodor]
b. Sp A. What does he (= John) think of Sam?
Sp B. Sam <u>he doesn't like</u> – think of someone else.

More recently, Fanselow and Lenertovà (forthcoming) arrive at a similar conclusion for extracted complements in German and Czech (root) declaratives. They show that left fronted complements in German declaratives can correspond to a narrow focus (18a), a topic (18b) or be part of a wider focus (18c).

(18) a. [What did you see there?]
Eine LaWIne haben wir gesehen!
(An avalanche-acc have we seen!)
b. [I've heard the mayor has been arrested. Who reported him to the police?]
Den Bürgenmeister hat wohl der Villenbesitzer angezeigt.
(The mayor.acc has supposedly the villa-owner.nom reported)

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

c. [What's new , What happened?]
Einen Hasen habe ich gefangen. (A rabbit.acc have I found)

They call the third case *sub-focus fronting* because only part of the focalized material is being fronted; their general conclusion, which indeed also applies to French, is that, contra what has been assumed in most generative analyses, leftwards movement is not triggered by informational features.

A similar observation has also been made for the extraction construction called "focalization" in Italian, in Godard and Marandin (2006, 2007). It has been proposed that the filler's content represents the focus (hence, the name, see e.g. Benincà et al. 1988). However, this is not the case, as is clearly shown by Godard and Marandin's examples (the observation is valid for positive as well as negative fillers, which is the main target of their study). The filler can be a narrow focus (19a), but it can also be included in an all focus sentence (19b) or be a partial, contrastive topic (20), thus included in the ground.[7]

(19) a. Sp. A. A chi ha parlato Maria per tutta la serata?
(To whom did Maria talk the whole evening?)
Sp. B. *A suo ex-ragazzo* ha parlato (Maria)
(She talked to her ex-boy friend)
b. [A mother opens the door of a room where there is a child. Something appears to be broken. The child says:]
Niente non ho fatto. (I have not done anything!)

(20) [A and B are colleagues in linguistics; they know that the exam had a syntactic and a semantic part]
Sp. A. I tuoi studenti hanno riuscito l'esame?
(Did your students do well at their exam)
Sp; B. *La sintassi* hanno capito, la semantica, invece, non c'è verso!
(They understood the syntax, but the semantics is a disaster)

There is a clear conclusion: there is no strict correlation between extraction (syntactic structure) and information structure. In other words, leftward movement is not "triggered" by information structure. With respect to the phenomenon under study, we can say that we have one construction, which diverges into two sub-constructions, depending on their information structure uses (the filler is a narrow focus, or the whole construction is all focus). One aspect remains to be studied in more detail: the phonological properties. Such declarative sentences with an NP filler have a specific prosody; the NP constitutes an intonational phrase (IP) in itself (Sabio 1995). However, it remains to be seen whether the prosody of the sentence is the same when the NP is a narrow focus, and when the sentence is all focus.

4. Illocutionary properties

The pragmatic properties are a crucial aspect of the constructiǹ. This is shown by a number of observations. First, the construction cannot be embedded: see the contrast between (21a) and (5a), (21b) and (9a). Second, the sentence can only be declarative, it cannot be

[7] A systematic comparison of the French and Italian constructions is still to be done; the comparison should also include the Spanish construction studied in Leonetti and Escandell-Vidal (forthcoming).

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

interrogative (22a) or injunctive (imperative, (22c)). However, it can be a declarative sentence with a questioning or injunctive value: see (9b), (22b) and (22d) (assuming that *Advil* is the name of a medication).

(21) a. *Je ne sais pas si [*dix-sept ans* il avait à l'époque]
I don't know whether eight years he had at the time
b. *On nous raconte [*qu'une antenne* ils lui ont jeté sur la tête]!
(we are told that an antenna they threw on his head)
(22) a. **Une antenne*, est-ce qu'ils lui ont jeté sur la tête?
(an antenna, did they throw on his head?)
b. *Une antenne* ils lui ont jeté sur la tête? (an antenna they threw on his head?)
c. **De l'advil*, prends! (Advil take!)
d. *De l'Advil*, tu prends, n'oublie pas! (Advil, you take, don't forget!)

These two properties show that the construction is defined as an utterance type, rather than being simply defined syntactically. We assume that sentence types are associated with a type of content (see e.g. Ginzburg and Sag 2000). The content of a declarative clause is a proposition, while that of an interrogative clause is a question, for instance. This is assumed in section 3.1 above, where we propose that the content of the construction is a (thetic) proposition. To account for the utterances in (22b,d), we say that a clause type can be associated with different speech act types; thus, a declarative clause can be associated not only with an assertion, but a confirmation query (22b), or a confirmation request (22d), see Beyssade and Marandin (2006).[8]

When used with a questioning value, the construction cannot be associated with a straightforward query, as is an interrogative clause: it is incompatible with a yes-no question (22a), as well as with a wh-question *(*Deux heures, qui attendrait ?* Two hours, who would wait-for); even when it has the form of a declarative, and looks as if were the equivalent of a yes-no question, it is biased. Thus, the continuation *ou pas* is inappropriate in (23b), while the continuation *n'est-ce pas* is appropriate (23c). This shows that the construction is associated with a confirmation query, with a bias towards a positive answer.

(23) a. *Huit ans* il avait à l'époque ? (8 years he had at the time?)
b. #*Huit ans* il avait à l'époque ou pas ? (8 years he had at the time or not?)
c. *Huit ans* il avait à l'époque, n'est-ce pas ? (8 years he had at the time, didn't he?)

When the construction is used as an answer to a question, the extracted NP can function as a narrow focus, as in (24) where it is an informational focus, but also in (25, 26) where it has a contrastive or correcting interpretation:

(24) Sp A. Comment ça s'appelait? (What was it called?)
Sp B. *L'écho de la Lys*, ça s'appelait. (*L'écho de la Lys* it was called)
(25) [the radio announces a new publication by AB]
Sp A. AB, la socialiste?(AB, the socialist?)
Sp B *Extrême-gauche*, elle était. (Extreme left, she was) [on the fly]

[8] Alternatively, we can say that a declarative clause is always associated with an assertion, (22b,d) representing indirect speech acts. The choice depends on a general discussion which is outside the limits of this study. We take "assertion", "query" and "request" to denote speech acts (as opposed to "proposition", "question" and "injunction", which denote contents).

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

(26) Sp. A. Tu as vu Dominique? (you have seen Dominique?)
Sp B. *Frédérique* elle s'appelle, pas Dominique (Frédérique she's called, not Dominique)

However, most often, it seems that it is not part of a smooth progression. It frequently contributes information that is relevant to the discourse topic (the question under discussion), rather than being a straightforward continuation (Ginzburg 2008). For instance, in (11), it solves the question of the name of the newspaper that was left unresolved in the preceding discourse, but this question itself is an aside rather than a piece of information contributing to the progression. In (14a), the construction is part of the story (it describes an event), but, as we have seen, at the same time, it stops the narrative sequence (see section 3.2.1 above). It is also frequently used as a means to correct the other speaker precisely on the point denoted by the NP, as in (25) and (26). Moreover, it is usually associated with speakers' psychological attitudes (surprise, admiration, disgust, justification, etc.).[9]

What role can we ascribe to the preposed NP? In a general way, it seems that the partition of the sentence in this construction serves to make the utterance more dramatic, highlighting that part of the situation that is particularly worth of notice, justifying a pause in the discourse, a reaction on the part of the audience etc. The role of the preposed NP then appears to be that of a "figure" (reminiscent of the figure-ground distinction in cognitive linguistics). For this reason, we call the whole construction a "dramatic extraction".

5. An HPSG analysis

We formalize our analysis within the framework of Head-driven Phrase structure grammar (Pollard and Sag 1987, 1994, Sag et al. 2003), a grammar model which makes use of the notion of construction, and which is well equipped for the study of interfaces.

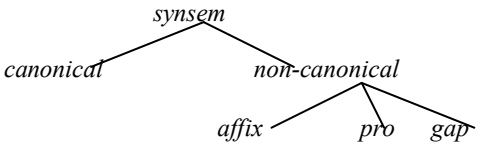
5.1 The Syntax of the extraction

The construction is a straightforward case of extraction, with a long distance dependency between the filler phrase and a predicate missing an argument (Ginzburg and Sag 2000). We represent the predicate missing an argument by typing one element in its argument structure as non canonical or gap. An element typed as *gap* belongs to the argument structure of a predicate, but is not locally realized (it does not match with a subject or a complement, or a specifier), unlike the arguments typed as *canonical* which are realized locally. The elements which belong to argument structures (bundles of syntactic (CAT) and semantic (CONT) features, called *synsem*) form the hierarchy in (27); only the canonical elements can be phrases; thus, the arguments which are not realized as phrases are affixes (we consider some pronominal clitics to be affixes), null pronouns and gaps. The constraint in (28) says that the argument structure of a word is made of the subject, the specifier and complements, to which one must add non-canonical elements (among which the gaps); "A", "B" and "C" note lists as well as "< x>" (including the empty list, which can be noted < >). A transitive verb whose object is gapped is described in (29). The notation "{ }" or S notes a set (the SLASH value is a set; in (29) the verb whose object is gapped may inherit other SLASH values besides the object gap, hence the union value).

[9] This may be characteristic of all focus sentences (as opposed to narrow focus ones). If this is true, it would further distinguish the two uses of the structure.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

(27) Hierarchy of synsem types



(28) Argument Realization Principle

$$word \Rightarrow \begin{bmatrix} \text{SYNSEM} \mid \text{LOC} \mid \text{VALENCE} \begin{bmatrix} \text{SUBJ} & \text{A} \\ \text{SPR} & \text{B} \\ \text{COMPS} & \text{C} \end{bmatrix} \\ \text{ARG-ST} \ (\text{A} + \text{B} + \text{C}) + \text{list}(non\text{-}canonical) \end{bmatrix}$$

(29) a verb with an extracted NP complement:

$$\begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{ARG-ST} & < canonical\,\text{NP-}i, \begin{bmatrix} gap \\ \text{LOCAL}\ [1] \\ \text{SLASH} & \{[1]\} \end{bmatrix} > \\ \text{VALENCE} \begin{bmatrix} \text{SUBJ} & < \text{NP-}i > \\ \text{COMPS} & < \ > \end{bmatrix} \end{bmatrix} \\ \text{SLASH} \ \{[1]\} \ \cup \text{S} \end{bmatrix} \end{bmatrix}$$

The type *gap* is described as in (30); it contains local features (the same syntactic and semantic features as in synsem, except that here is no information concerning long distance dependencies), and a SLASH feature, whose value is identical to these local features, but which allows them to be treated non locally. The SLASH feature is inherited from the predicate by the phrase it is the head of, following the generalized head feature principle (31a), which says that the features are identical on the phrases and the head of the phrase, unless otherwise specified by a specific construction (the notation '/' indicates a default constraint). The SLASH feature is propagated from that phrase to the higher predicate of which the phrase is an argument, until the top of the construction is reached (see the SLASH amalgamation constraint in (31b)).

(30) $$gap \Rightarrow \begin{bmatrix} \text{SYNSEM} & [1] \\ \text{SLASH} & \{[1]\} \end{bmatrix}$$

(31) a. Generalized Head Feature Principle:

headed-phrase [SYNSEM /[1]] => HEAD-DAUGHTER [SYNSEM / [1]]

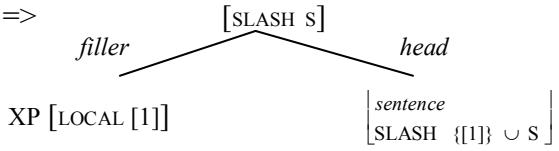
b. SLASH Amalgamation Constraint:

$$word \Rightarrow \begin{bmatrix} \text{SYNSEM} & \text{S1} \cup \ldots \cup \text{Sn} \\ \text{ARG-ST} & < [\text{SLASH S1}, \ldots, [\text{SLASH Sn}] > \end{bmatrix}$$

The top of the construction is a filler-head phrase, which stops the propagation of the SLASH feature, when the syntactic and semantic features of the filler daughter are identical to those mentioned in the SLASH.

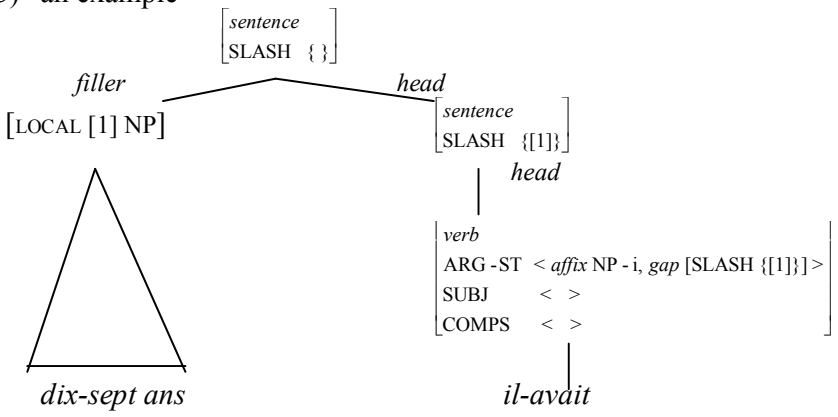
BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

(32) *head filler phrase* =>



The syntactic structure of the sentence *dix-sept ans il avait* is thus as follows:[10]

(33) an example



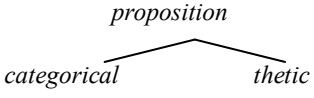
5.2 The hierarchy of clause types

We leave aside the question of the semantic combination of the head sentence with the extracted NP (which can be done by adopting an HPSG version of lambda application, along the lines proposed in Webelhuth 2007, for instance), and we concentrate on what constitutes the originality of the construction, that is, the clause type. As we have seen, only a declarative clause is acceptable, with several specific properties:
(i) its content is a thetic proposition;
(ii) it is a main clause (it cannot be non-embedded);
(iii) it contains a phrase which functions either as a narrow focus or as a 'figure' (as opposed to the ground, in cognitive linguistic terms). We call it a 'center' for the clause, to generalize over the two possibilities.

The phrase types (among which the filler- head phrase, (32)) are one dimension of the analysis of expressions. Clauses constitute a different dimension. Clauses are essentially defined by the type of content (Ginzburg and Sag 2000). The content of declarative clauses is a proposition. We distinguish two subtypes of propositions, thetic and categorical.

(34) *declarative clause* => [CONT *proposition*]

(35)



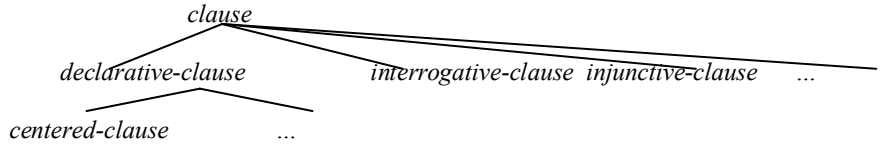
To account for the presence of a phrase having the role of a center, we set up a subtype of declarative clause, the centered clause, which could also be used for clefted constructions,

[10] Although it is a totally independent issue, we follow here Miller and Sag (1997) analysis of pronominal clitics as affixes; hence the prefix analysis of the clitic *il*. In formal registers, subject clitics may remain independent words, but it is not the case in the informal register characteristic of the construction studied here.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

with a feature CENTER, whose value is the content of a phrase.[11] In the case of the dramatic extraction construction, its value is the content of the filler. Our construction also has a feature [Independent Clause +, or IC +], indicating that it cannot be embedded.

(36)



5.3 The dramatic extraction construction

To sum up, the dramatic extraction construction inherits from the head-filler-phrase and the centered-clause:

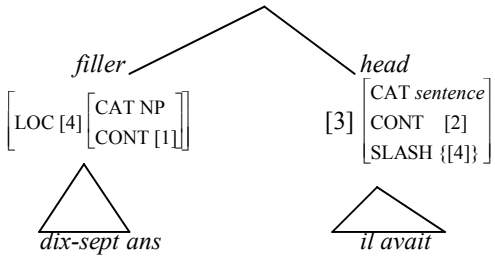
(37) *dramatic-extraction-construction* =>

$$\begin{bmatrix} \textit{head-filler-phrase} \,\&\, \textit{centered-clause} \\ \text{CAT} \quad [\text{IC} +] \\ \text{CONT} \quad \begin{bmatrix} \textit{thetic - proposition} \\ \text{SOA } [2] \end{bmatrix} \\ \text{CENTER } [1] \\ \text{HEAD-DAUGHTER } [3] \begin{bmatrix} \text{CONT } [2] \\ \text{SLASH } \{[4]\} \end{bmatrix} \\ \text{DAUGHTERS } \langle [\text{LOCAL } [4] \, [\text{CONT } [1]], [3] \rangle \end{bmatrix}$$

An example integrating the different aspects of the analysis is given in (38).

(38) a more complete example

$$\begin{bmatrix} \textit{centered clause} \,\&\, \textit{filler-head-phrase} \\ \text{CAT} \quad \textit{sentence} \\ \text{IC} \quad + \\ \text{CONT} \quad \begin{bmatrix} \textit{thetic proposition} \\ \text{SOA } [2] \end{bmatrix} \\ \text{CENTER } [1] \\ \text{SLASH } \{\,\} \\ \text{HEAD-DAUGHTER } [3] \begin{bmatrix} \text{CONT } [2] \\ \text{SLASH } \{[4]\} \end{bmatrix} \\ \text{DAUGHTERS} \quad \langle [\text{LOC } [4] \, [\text{CONT } [1]], [3] \rangle \end{bmatrix}$$



[11] Cleft sentences in French have been shown, informationally, either to contain a narrow focus (the content of the clefted element) or to constitute an all focus sentence (*Qu'est-ce qui se passe ? C'est le téléphone qui sonne* ! 'what happens ? it's the phone that's ringing'). See Doetjes et al. (2004).

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

6. Conclusions

There exists in spoken French a construction consisting of a declarative clause with an initial, extracted constituents. All types of complements are eligible for that construction (even predicative or non referential ones), as well as specifiers of nouns, but indefinites are preferred. This French construction seems to correspond to the construction sometimes called 'Focalization' in other languages, such as Italian. However, we show that the preposed complement is not always associated with the same informational status: it can be interpreted as a narrow focus, but it can also be part of an all-focus utterance. More generally, the conclusion is that there is no one-to-one mapping between syntactic structure and information structure, and a more modular grammar architecture is necessary. In addition, we show that the preposed constituent cannot be interpreted as a sentence Theme, in a categorical proposition, and that the construction always denotes a thetic proposition. More generally, there is no one-to-one mapping between semantic structure and information structure (a thetic proposition does not always correspond to an all-focus utterance).

This study, which is formalized in the framework of HPSG grammar, shows the need for constructions, that is, linguistic objects which constrain the association between syntax, semantics and discourse properties.

Anne Abeillé
University Paris 7
LLF
anne.abeille@linguist.jussieu.fr

Danièle Godard
CNRS, LLF and University Paris 7
daniele.godard@linguist.jussieu.fr

Frédéric Sabio
University of Provence
LPL
frederic.sabio@orange.fr

References

Abeillé A., Bonami, O., Godard, D., Tseng, J. 2004. The syntax of French de-N' phrases. In S. Müller (ed.), *Proceedings of the HPSG Conference*, 6-26. Stanford: CSLI on line Publications, <http://csli-publications.stanford.edu>.

Abeillé, A, Godard, D., Sabio, F. 2008. Two types of NP preposing in French. In S. Müller (ed), *Proceedings of the 15th HPSG Conference*, 306-324. Stanford: CSLI on-line Publications, <http://csli-publications. stanford.edu>.

Benincà, P., Salvi, G., Frison, L. 1988. L'ordine degli elementi della frase e le costruzioni marcate. In L. Renzi (ed.), *Grande grammatica di consultazione*, vol. 1, 115-215. Bologna: Il Mulino.

Beyssade, C. and Marandin, J.-M. 2006. From complex to simple speech acts: A bidimensional analysis of illocutionary. Paper presented at Brandial 2006, University of Potsdam.

Blanche-Benveniste, C., Rouget, C., Sabio, F. (eds.). 2002. *Choix de textes de français parlé*. 36 extraits. Paris: Champion, Collection "Les fançais parlés").

Büring, D. 1997. *The 59th StreetBridgeAaccent The Meaning of Topic and Focus*. London: Routledge.

Doetjes, J., Rebuschi, G., Rialland, A. 2004. Clefted sentences. In C. Corblin, H. de Swart (eds.), *Handbook of French Semantics*, 529-552. Stanford: CSLI Publications.

Fanselow G., Lenertovà, D. forthcoming. Left peripheral focus: Mismatches between syntax and information structure. *Natural Language and Linguistic Theory*.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)

Ginzburg, J. 2008. *Semantics and Conversation*. Stanford: CSLI Publications.
Ginzburg, J., Sag, I. A. 2000. *Interrogative Investigations*. Stanford : CSLI Publications and Chicago: The University of Chicago Press.
Godard, D., Marandin, J-M. 2006. Reinforcing negation: The case of Italian. In S. Müller (ed.), *Proceedings of HPSG 06*, 174-194. Stanford: CSLI Publications.
Godard, D., Marandin, J-M. 2007. Aspects pragmatiques de la négation renforcée en italien. In F. Floricic (ed.), *La négation dans les langues romanes*, 137-160. Amsterdam/Philadelphia: John Benjamins.
Jacobs, J. 2001. The dimensions of topic-comment. *Linguistics* 39: 641-681.
Kim, Y. 1998. Information articulation and truth conditions of existential sentences. *Language and Information* 1: 67-105.
Kratzer, A. 1995. Stage level and individual level predicates. In G. Carlson and F. J. Pelletier (eds.), *The Generic Book*, 125-175. Chicago: The University of Chicago Press.
Ladusaw, W. 1994. Thetic and categorical, stage and individual, weak and strong. In M. Harvey and L. Santelman (eds.), *Proceedings of SALT IV*, 220-229. Ithaca, NY: Cornell University Press.
Leonetti, M., Escandell-Vidal, V. forthcoming. Fronting and verum focus in Romance. In A. Dufter and D. Jacob (eds.), *Focus and Background in Romance Languages*, Amsterdam/Philadelphia: John Benjamins.
Marandin, J-M. 2007. Thème de phrase. In D. Godard, F. Corblin, L. Roussarie (eds.), *Dictionnaire de sémantique en ligne*. <http://www.semantique- gdr.net/dico/index.php/>.
Miller, P. H. And Sag, I. A. 1997. *French Clitic Movement without Clitics or Movement*. Stanford: CSLI Publications.
Milner, J-C. 1978. *De la syntaxe à l'interprétation: quantités, insultes, exclamations*. Paris: Le Seuil.
Pohl, J. 1984. Documents pour servir à l'étude des phrases du type *Les fleurs, j'aime*. *Romanistisches Jahrbuch* 35: 36-58.
Pollard, C. and Sag, I. A. 1987. *Information-based Syntax and Semantics*. Stanford: CSLI Lecture Notes (No. 13).
Pollard, C. and Sag, I. A. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Prince, E. 1981, Toward a taxonomy of given-new information. In P. Cole (ed.), *Radical pragmatics,* 223-255. New York: Academic Press.
Prince, E. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In P. Culicover and L. McNally (eds.), *The Limits of Syntax*, 281-302. New York: Academic Press.
Prince, E. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. <ftp://babel.ling.upenn.edu/papers/faculty/ellen_prince/texas.ps>.
Roberts, C. 2004. Context in dynamic interpretation. In L. Horn and G. Ward (eds.), *Handbook of Contemporary Pragmatic Theory*, 197-220. Oxford: Blackwell.
Sabio, F. 1995. Micro-syntaxe et macro-syntaxe: l'exemple des compléments antéposés en français. *Recherches sur le français parlé* 13: 111-155.
Sabio, F. 2006. L'antéposition des compléments en français contemporain: l'exemple des objets directs, *Linguisticae Investigationes* 29: 173-182.
Sag, I. A., Wasow, T. and Bender, E. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications.
Webelhuth, G. 2007. Complex topic comment structures in HPSG. In S. Müller (ed.), *Proceedings of the HPSG 07 Conference*, 306-322. Stanford: CSLI on-line Publications, <http://csli-publications.stanford.edu>.

BDD-A9808 © 2009 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.42 (2026-11-15 20:55:33 UTC)