
Constituency and word order in French subject inversion

OLIVIER BONAMI, DANIELE GODARD AND JEAN-MARIE MARANDIN

In Gosse Bouma, Erhard Hinrichs, Geert-Jan M. Kruijff & Richard Oehrle (eds). 1999. *Constraints and Resources in Natural Language Syntax and Semantics*, pp. 21–40. Stanford, CA: CSLI Publications.

1.1 Introduction

Subject NP inversion is a very common phenomenon in French.¹ It comes in three varieties, which are not clearly distinguished in the linguistic tradition although they have different properties: (i) inversion in extraction contexts (1); (ii) heavy subject NP inversion (2); (iii) inversion in spatio-temporally dependent clauses, instantiated in three contexts : time adverbials (3a), subjunctive complements (3b), and sentences with athetic interpretation in a narrative (3c).

- (1) Voici le texte qu'a écrit **Paul**.
'Here is the text that Paul wrote.'
- (2) Ont accepté notre proposition **les députés de la majorité ainsi que les non-inscrits**.
'The MPs of the majority as well as the nonregistered ones have accepted our proposal.'
- (3) a. Dès que se lève **le soleil**, le coq chante.
'As soon as the sun rises, the rooster sings.'
 b. Je veux que soit invitée **Marie**.
'I want that Marie be invited'
 c. (Alors) arriva **Marie**.
'(Then,) Marie arrived.'

¹We call 'subject NP inversion' what has been called 'stylistic inversion' since Kayne 1972. Note that it differs sharply from clitic subject inversion as shown there. Throughout this paper, inverted subjects are typeset in boldface.

In this paper, we restrict our attention to the first variety, extraction-triggered inversion (ETI). It has been known since (Kayne and Pollock 1978, Milner 1978, Zaenen 1983) that inversion is allowed along extraction pathways. We present a novel analysis of the phenomenon which is couched in a linearization-based HPSG framework and crucially relies on new data concerning inversion in embedded clauses.² Essentially, the analysis is as follows: the constituent structure of the inverted sentence is identical to that of the usual non inverted sentence (an NP VP construction) and the position of the subject NP is due to the union of the VP word order domain into that of the sentence. This analysis easily extends to another case, locative inversion, which has not been recognized as an instance of ETI in previous studies.

1.2 Properties of inverted subjects

Subject inversion is possible in all well-known extraction contexts in French: relatives, *wh*-interrogatives or exclamatives, clefts, PP topicalizations. In this section, we illustrate the general properties of ETI with relative clauses.

Preverbal subjects in French contrast with postverbal objects in a number of ways. Since the inverted NP in extraction contexts shares properties with both, the question of its grammatical function is not trivial. The relevant observations are summarized in figure 1 (See Marandin 1997 and references therein).

Like noninverted subjects and unlike objects, the inverted NP can bind the anaphor *se* (4). Objects but not subjects allow the quantitative *en* construction; the inverted NP patterns with subjects (5). The same bare Qs that can be objects can be inverted subjects (6). The floating *Q...de N* construction, which is available for objects but not for subjects, is impossible for the inverted NP (7). Like objects but unlike noninverted subjects, the inverted NP can take the form *de N* in negative contexts (8). The interrogative determiner *combien*, which can be extracted out of an object but not out of a noninverted subject, can be extracted from the inverted NP (9). And finally, while objects never agree with the finite verb in French, and preverbal subjects agree both in number and person, the inverted NP only agrees in number with the verb (10).³

²We leave aside the information structure and discourse properties which have been the main focus of the grammatical tradition (e.g., Le Bidois 1950). See Lambrecht and Polinsky 1997 for a crosslinguistic analysis and Bonami et al. 1998 which shows that subjects in extraction-context inversions cannot be topics in French.

³This difference between number and person agreement is not specific to NP inversion. Similarly, the infinitival complement in causative constructions does not

- (4) le miroir où se_i voit **Paul**_i
the mirror where SE sees Paul
‘the mirror where Paul sees himself’
- (5)*les livres qu’ en ont lu **trois**
the books that EN have read three
‘the books which three of them read’
- (6) a. Certains/quelques-uns viendront.
‘Some (of them) will come.’
b. Paul connaît certains/*quelques-uns.
‘Paul knows some (of them).’
c. un problème que connaissent **certains** /***quelques-uns**
a problem that know some some
‘a problem that some (of them) know’
- (7)*l’ année où sont beaucoup parus **de best-sellers**
the year where are many published DE best-sellers
‘the year during which many best sellers were published’
- (8) une maison où ne viennent plus jamais **d’ enfants**
a house where NE come no-more never DE children
‘a house where kids do not come anymore’
- (9) Combien sont venus **de clients** aujourd’hui?
How-many are come DE clients today
‘How many clients came today?’
- (10) l’ immeuble où habitaient /*habitions **Marie et moi**
the building where lived-3.PL lived-1.PL Marie and I
‘the building where Marie and I lived’

In order to make sense of this complex pattern, we propose to rely on the case/function distinction: the inverted NP can be analyzed either as an accusative subject or as a nominative object. While most of the properties in figure 1 are compatible with either analysis, the restriction on bare quantifiers unexpectedly favors a subject analysis of the inverted NP. The behavior of bare Qs in general points to a constraint on case rather than on grammatical function.

The distribution of bare Qs is extremely intricate. They divide into two groups: (a) those which can be objects: *beaucoup* ‘a lot’ (in its non-human and non-anaphoric use: *Il apportera beaucoup en travaillant au projet* ‘he will contribute a lot by working on the project’),

agree in person with a postverbal causee (*On a fait se/*nous lever tôt mon frère et moi* ‘They made my brother and me wake up early’).

	inverted NP	subject	object
1. binding of <i>se</i>	yes	yes	no
2. quantitative <i>en</i> construction	no	no	yes
3. bare Qs <i>tous, beaucoup</i>	no	yes	no
4. floating <i>beaucoup (... de N)</i>	no	no	yes
5. <i>de N</i> in negative context	yes	no	yes
6. <i>combien</i> extraction	yes	no	yes
7. number agreement with finite V	yes	yes	no
8. person agreement with finite V	no	yes	no

FIGURE 1 Properties of subjects, objects and inverted NPs in ETI

*certain*s ‘some’ and *chacun* ‘each’ (which is human or anaphoric: *Dieu examine chacun avec indulgence* ‘God examines each person leniently’), *tout* ‘everything’ and *rien* ‘nothing’ (which are non-human and non-anaphoric); (b) those which cannot be objects: *beaucoup* ‘many’ (if anaphoric or with human reference), cardinals, *quelques-uns* ‘a few’, *tous* ‘all’. Neither group forms a natural class with respect to quantification, anaphoric/non-anaphoric use or human/non-human reference. The restrictions must be encoded on the lexical items themselves: some bare Qs simply fail to have an accusative form (Abeillé and Godard 1998).

Going back to ETI, the crucial observation is that those bare Qs that can be inverted NPs are precisely those which can be objects (6). This follows if inverted NPs are accusative, and is problematic if they are nominative. However, inverted NPs cannot be accusative complements, or they would not differ from objects. Thus, inverted NPs are accusative subjects (see also Abeillé 1997).

More precisely, if inverted NPs in ETI were accusative complements, they would only differ from objects in their position on ARG-ST. Such an analysis does not allow for a straightforward treatment of the constraints on quantitative *en*: on this approach, the ungrammaticality of (5) would result from a constraint stating that *en* cannot be linked to the first member of the ARG-ST. But this cannot be correct, since in other constructions, we do find *en* linked to the first member of the ARG-ST. It is the case in some instances of variety (iii) inversion (see paragraph 1.1).

- (11) a. Entrèrent trois hommes.
 ‘Three men came in.’
 b. En entrèrent trois.
 EN came-in three
 ‘Three of them came in.’

Since the postverbal NP in (11a) has properties untypical of objects (e.g.

it agrees un number with the verb), we need a four way rather than a three way distinction: between noninverted subject NPs, inverted NPs in ETI, postverbal NPs in sentences such as (11), and objects. In our analysis, the first are nominative subjects (and the first member of the ARG-ST), the second accusative subjects (and the first member of the ARG-ST), the third accusative complements which are the first member of the ARG-ST⁴ and the fourth accusative complements which are not the first member of the ARG-ST (that is, objects). The quantitative *en* construction is restricted to complements.⁵

Given the analysis of the inverted NP in ETI as an accusative subject, we can account for the other properties of the inverted NP in the following way. (i) All and only subjects can bind the anaphor *se*. (ii) The form *de N* is always accusative, a constraint which accounts for observation 5 and 6 in figure 1. (iii) Floating *beaucoup* can only be linked to an object. (iv) Finally, agreement data do not correlate with a subject/object contrast; rather they suggest that the verb always agrees in number with the first item on ARG-ST, but only agrees in person with a subject preceding it.

1.3 Inversion and phrase structure

1.3.1 A problem of word order

It has long been observed that the inverted subject can be linearized not only after the VP, but also between the verb and one of its complements:

- (12) a. la lettre qu' enverra à la direction **le patron**
 the letter that send-FUT to the management the boss
 'the letter that the boss will send to the management'
 b. la lettre qu'enverra **le patron** à la direction

This fact can be accounted for in two different ways. ETI sentences can be assumed to have a flat structure, the inverted subject being a sister of the lexical verb and of its complements.⁶ Since the order between NP and PP is unconstrained in French (modulo discursive factors), the two orders in (12) would be accounted for directly. Alternatively, one may assume that either the NP or the PP is extraposed.⁷

⁴This is the analysis we adopt for what is essentially an unaccusative realization of verbs.

⁵This analysis correctly predicts that predicative NP complements allow the quantitative *en* construction (*Un chef, il ne veut pas en devenir un* 'He does not want to become a leader').

⁶Such an analysis requires positing a new phrase type *head-subject-complements*.

⁷This is the standard hypothesis in transformational analyses. Depending on the specific proposal, the subject NP is assumed to be either right adjoined to the VP, or in its base position, which is to the left or to the right of the VP (see de Wind

However, none of these analyses can account for the word order illustrated in (13), a piece of data which has not been brought to light in previous studies. When the embedded V[*inf*] is slashed, the higher subject may occur between the embedded V and its complements. Note the potentially unbounded distance between the main verb and the VP among whose constituents the inverted subject is found.

- (13) a. le livre que pouvait recommander **le patron du**
 the book that could recommend the head of-the
labo à cet étudiant
 lab to this student
 ‘the book that the head of the lab could recommend to this student’
- b. le livre que croyait pouvoir recommander **le patron**
 the book that thought can recommend the head
du labo à cet étudiant
 of-the lab to this student
 ‘the book that the head of the lab thought he could recommend to this student’

1.3.2 A problem for extraposition

An extraposition analysis for (13) would have to assume that extraposition is unbounded. Figure 2 illustrates this fact: the PP has to cross two VPs (or S) boundaries. The dashed lines indicate different conceivable attachments for the inverted subject. Note that generating the subject in a different ‘base’ position does not solve the problem since the PP has to be extraposed in any case to yield the right order.

But extraposition is clearly clause-bounded in French. Example (14) illustrates the well-known fact that extraposition of a relative clause out of an NP is clause-bounded.

- (14) a. Je demanderai [à rencontrer [des lecteurs qui ont aimé
 I ask-FUT À meet DES readers who have liked
 mon livre]] à mon éditeur.
 my book to my publisher
 ‘I will ask my publisher to meet readers who have liked my book.’
- b.*Je demanderai à rencontrer des lecteurs à mon éditeur qui ont aimé mon livre.

An extraposition analysis of (13) would involve a hypothetical PP extraposition. Example (15) shows that the purported PP extraposition would similarly have to be clause-bounded. Accordingly, the extraposi-

1995 and references therein for discussion).

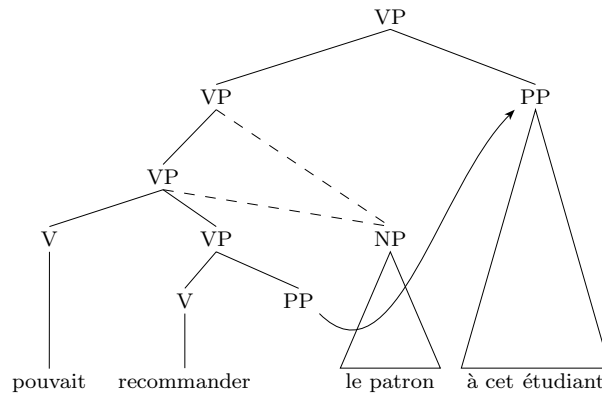


FIGURE 2 Extraposition analysis for (13a)

tion analysis cannot account for the contrast between the acceptability of (13) and the unacceptability of (15b).

- (15) a. Le patron du labo [disait [travailler sur ce sujet]
 the head of-the lab said work on this subject
 [à ses collaborateurs]].
 to his collaborators
 ‘The head of the lab said he worked on this subject to his
 collaborators.’
 b.*Le patron du labo disait travailler à ses collaborateurs sur ce
 sujet.

1.3.3 Problems for flat structures

Another possibility would be to assume a flattened structure for the main VP in (13), as in figure 3:⁸ in the spirit of e.g., Hinrichs and Nakazawa 1994, we could assume that the verb whose subject is inverted is a composition verb inheriting the arguments of its infinitival V complement. This would create a flat structure with the inverted subject at the same level as the V[*inf*], thus allowing the unusual position of the inverted subject. Such an analysis encounters three serious empirical problems.

Non-inherited arguments. First, it is not clear what type of arguments would be inherited by the verb whose subject is inverted. In causative constructions, which have been argued to instantiate a flat complementation (e.g., Abeillé et al. 1998), clitics are inherited by the

⁸The flat structure seems to be the only purely phrase-structural option: to account for the data in (13) with a hierarchical complementation, we would have to postulate a new type of ‘upside down’ unbounded dependency where the subject of the main verb would be at the same level as the complements of the V[*inf*].

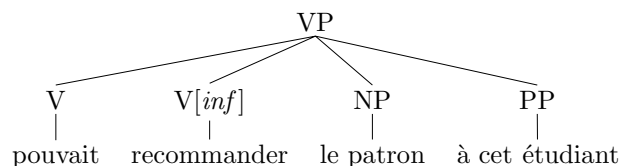


FIGURE 3 Flat structure analysis for (13a)

head verb which is their host (16). This is not the case in ETI (17), since the clitic argument only occurs on the downstairs verb.

- (16) a. Paul le fera lire aux étudiants.
 Paul it make-FUT read to-the students
 ‘Paul will make the students read it.’
 b. *Paul fera le lire aux étudiants.
- (17) a. *un message que lui veut envoyer **Paul**
 a message which to-him wants send Paul
 ‘a message which Paul wants to send to him’
 b. un message que veut lui envoyer **Paul**

Hence, the purported composition verb in ETI would inherit only the canonical complements from its $V[inf]$ complement. However, this turns out not to be the correct generalization.

The floating *Q beaucoup* is inherited by composition verbs (auxiliaries and causative verbs) as shown by its occurrence between the auxiliary and the past participle *fait* (18a). The inverted NP is perfectly compatible with such inheritance (18b).

- (18) a. les étudiants auxquels le prof de math en a
 the students to-whom the teacher of math EN has
 beaucoup fait faire
 many made do
 ‘the students whom the math teacher made do many (exercises)’
 b. les étudiants auxquels en a beaucoup fait faire **le prof de math**

On the other hand, *beaucoup* cannot climb out of a VP whose head is not a composition verb. Subject inversion does not modify this behavior.

- (19) a. *les étudiants auxquels le prof de math a beaucoup
 the students to-whom the teacher of math has many
 voulu en donner
 wanted EN give
 ‘the students to whom the math teacher wanted to give many (exercises)’

- b.*les étudiants auxquels a beaucoup voulu en donner **le prof de math**

Accordingly, the putative argument-composition *vouloir* in (19b) does not inherit *beaucoup*, although *beaucoup* is canonical. The flat structure analysis would have to prevent not only clitics but also canonical bare Qs from being inherited. Such a stipulation would be unfortunate, since clitics and bare Qs are precisely the items whose distribution gives a strong argument in favor of a flat structure for causatives and auxiliaries; thus the putative composition verbs in ETI would be unable to inherit any of the items for which there is independent evidence of inheritance.

Binding. The binding data present a second challenge for argument composition. The expression *l'un...l'autre* is an anaphor which must be bound by an o-commanding argument in every ARG-ST list where it occurs, as shown by its behavior in causatives. In (20a), both arguments of the V[*inf*] (*des patients, l'un après l'autre*) are inherited by *faire*; the anaphor is bound by its antecedent on both the ARG-ST of *s'occuper* and *faire*. On the other hand, in (20b), the cliticized antecedent *en* is 'trapped' on the downstairs verb and fails to occur on the ARG-ST of *faire*.⁹ Thus *l'un après l'autre* is not bound on the main verb's ARG-ST, which results in agrammaticality.

- (20) a. Sa rigueur le fera s' occuper [des patients]_i
 His orderliness him make-FUT SE take-care of-the patients
 [l' un après l' autre]_i.
 the one after the other
 'His orderliness will make him take care of the patients one after the other.'
- b.*Sa rigueur le fera s' en_i occuper
 His orderliness him make-FUT SE EN take-care
 [l' un après l' autre]_i.
 the one after the other
 'His orderliness will make him take care of them one after the other.'

If NP inversion were due to composition, we would expect (21b) to be similarly unacceptable: the cliticized antecedent (*les*) of the anaphor *l'un...l'autre* only belongs to the ARG-ST of *mélanger*. If the anaphor were inherited by the higher verb, it would be unbound. But (21b) is grammatical; thus, *l'un...l'autre* should not be inherited in ETI sentences.

⁹We follow Abeillé *et al.*'s 1998 analysis where composition-*faire* can take a complement hosting clitics if one of them is an inherent clitic such as the reflexive in *s'occuper*.

- (21) a. On a perdu le tube dans lequel notre préparateur
 We have lost the tube in which our assistant
 avait pensé les_i mélanger [les unes avec les autres]_i.
 had planned them mix the ones with the others
 ‘We have lost the tube in which the assistant had planned to
 mix them together (the colors).’
- b. On a perdu le tube dans lequel avait pensé les_i mélanger **notre préparateur** [les unes avec les autres]_i.

VP marking. The third problem for the composition analysis resides in the combined properties of adverbs and the marking system of infinitival VPs (Abeillé and Godard, 1997). The bare quantifier *tout* adjoins to the left of the V[*inf*] (Abeillé and Godard, 1998). On the other hand, an adverb such as *fréquemment* adjoins to the left of VP[*inf*] rather than V[*inf*]; this is shown by the fact that it cannot occur between *tout* and the verb (22).¹⁰

- (22) a. Il se résigne à fréquemment tout expliquer à ses élèves.
 He SE accepts à frequently all explain to his students
 ‘He accepts to explain frequently everything to his students.’
- b.*Il se résigne à tout fréquemment expliquer à ses élèves.

Certain verbs subcategorize for a VP ‘marked’ with *de* or *à*. That such elements combine with a VP is shown by their occurrence to the left of *fréquemment* (23a). We analyze them as phrasal affixes realized on the leftmost branch of the VP, in the spirit of Miller 1992. Crucially, subject inversion does not affect their position (23b).

- (23) a. C’ est un problème que notre médiéviste se résigne à
 It is a problem that our medievalist SE accepts à
 fréquemment expliquer aux étudiants de première année.
 frequently explain to-the students of first year
 ‘It’s a problem that our medievalist accepts to explain frequently to first year students.’
- b. C’est un problème que se résigne à fréquemment expliquer **notre médiéviste** aux étudiants de première année.

These data raise difficulty for a flat structure analysis because the notion of a VP leftmost branch disappears at the same time as the VP disappears. Without a VP constituent, it is impossible to account for the fact that the affix *à* occurs on the leftmost adjunct, if there is one,

¹⁰Note that this is not a scope problem: *Paul a tout expliqué à ses étudiants fréquemment* can mean that for each thing Paul explained to his students, he explained it frequently.

and on the V otherwise.¹¹ On the other hand, the data follow from the independently motivated analysis of *de* and *à* provided the main verb in ETI constructions takes a VP complement as it does in noninverted sentences.

We thus conclude that argument composition is an inadequate treatment for long-distance inversion.

1.4 A domain-union analysis

In our approach, the constituent structure of an ETI sentence is not different from that of a noninverted sentence. Subject inversion is the consequence of a different word order domain configuration. Specifically, we adopt Reape's (1994) domain union framework, and we assume that in inverted sentences, the slashed VP's word order domain gets unioned into the S domain. Having presented the data in more detail, we propose a treatment involving verb subtypes which we connect with domain union, and we give the major relevant linear precedence statements.

1.4.1 More data

Inversion is licensed by a gap in an NP.

- (24) a. Combien sont venus **de clients** ?
 How-many are come DE clients
 'How many clients came?'
- b. une brosse à habits dont se recourbait **la poignée**
 a brush to clothes of-which SE bent the handle
 'a clothes brush the handle of which bended'

It is never completely acceptable with verbs taking a complement sentence, whether the gap is local or inherited (*contra* Kayne and Pollock 1978); the complement sentence itself can of course be inverted (26).

- (25) a.*l' étudiant à qui disait **Marie** qu' elle ne viendrait pas
 the student to who said Marie that she NE come not
 'the student to whom Marie said that she would not come'
- b.*le livre qu' avait cru que Jean écrivait
 the book that had thought that Jean write-COND
mon éditeur parisien
 my publisher Parisian
 'the book that my Parisian publisher thought that Jean would write'

¹¹A similar argument can be made if *de* and *à* are analyzed as words: they would have to select a VP, but there is no VP for *à* to select in (23b) on a flat structure analysis.

- (26) le livre que mon éditeur avait cru qu'écrirait **Jean**
 'the book my publisher believed that Jean would write'

Subject inversion is always acceptable with verbs taking a VP[*inf*] complement: inversion with a subject control verb is illustrated in (27a), inversion with an object control verb is illustrated in (27b).

- (27) a. le livre que voulait offrir à ma fille **un libraire**
 the book that wanted offer to my daughter a bookseller
 'the book that a bookseller wanted to offer to my daughter'
- b. le livre que m' a convaincu d' offrir à ma fille
 the book that me has convinced DE offer to my daughter
un libraire
 a bookseller
 'the book that a bookseller convinced me to offer to my daughter'

However, only subject control or raising verbs allow the problematic order discussed above in 1.3.1 and illustrated again in (28a): the inverted subject must follow the embedded complements when the higher verb is an object control verb (28b).

- (28) a. le rôle bénéfique que lui semblait jouer **Pierre**
 the role favorable that to-him seemed play Pierre
 dans ce travail
 in this work
 'the favorable role that Pierre seemed to play in this work to him'
- b.*le livre que m' a convaincu d' offrir **mon libraire**
 the book that me has convinced DE offer my bookseller
 à Marie
 to Marie
 'the book that my bookseller convinced me to offer to Marie'

We analyze these data in the following way: the sequence of the embedded V, the higher subject and the complement of the embedded V indicates that all these items occur at the same level in the sentence's word order domain. The slashed VP complement of a subject control verb, but not that of an object control verb, may contribute its daughters to a higher word order domain.

1.4.2 The treatment

Verb types. We distinguish between two HEAD *verb* values: *non-inv-vb* (noninverted verb) and *extr-inv-vb* (extraction inverted verb). While we do not constrain the former (e.g. its slash value may be empty or non-empty), we constrain the latter to have a non-empty slash value, a

(canonical) accusative subject and no finite sentence complement. The following constraint holds for words.¹²

$$(29) \left[\begin{array}{l} \text{word} \\ \text{HEAD} \quad \text{extr-inv-verb} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{SUBJ} \quad \langle \text{NP}[\text{acc}] \rangle \\ \text{COMPS} \quad \text{list}([\text{HEAD} \text{ nonfin}]) \\ \text{SLASH} \quad \text{nelist} \end{array} \right]$$

We rely on a lexicalized treatment of extraction (Sag 1997), whereby heads inherit the SLASH value of their arguments and pass it to their mother. Thus (29) allows subject inversion to occur whether the gap is local or inherited, while correctly excluding (25). To account for the contrast between subject and object control verbs (28), we force the VP complement of the latter to be *non-inv-vb* and leave the complement of the former unspecified. The distinction between the verb subtypes is available for subcategorization since it is a HEAD value. Object control verbs are as follows:

$$(30) \left[\text{ARG-ST} \quad \langle \text{NP}, \text{NP}_i \rangle, \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{non-inv-verb} \\ \text{VFORM} \quad \text{inf} \end{array} \right] \\ \text{SUBJ} \quad \langle \text{NP}_i \rangle \end{array} \right] \right]$$

Domain Union. Following Reape 1994, we assume that every phrase is associated with a word order domain (DOM), which is the locus of LP statements. DOM is a list of signs, and the PHON value of a phrase is obtained by concatenating the PHON values of the members of its DOM list, respecting order. A binary feature UN(ion) constrains the relation between phrase structure and order domains: a [UN−] phrase is inserted in its mother’s domain as one item; on the other hand, the different items on a [UN+] phrase’s DOM value are inserted into the mother’s DOM instead of the phrase itself. In other words, the DOM value of a phrase is a list made of the [UN−] daughters and the elements of the DOM value of the [UN+] daughters.¹³

¹²We assume that the *finite/nonfinite* distinction applies uniformly to all head values. *list*([F *v*]) abbreviates a list of objects which all have the specification [F *v*].

¹³‘ \circ ’ notes the nondeterministic *shuffle* operation, which takes two lists and outputs a new list consisting of all their elements, preserving the order in the original lists. The DTRS feature, which is assumed here for brevity, shuffles the HD-DTR among the NHD-DTRS. (31) differs slightly from Reape’s domain principle in allowing head daughters to union; this is crucial to account for simple inversion sentences, where a head VP is unioned. Note that the data on VP marking in section 1.3.3 shows that both phrase structure and order domains are necessary for French; thus we do not follow Kathol 1995, which dispenses completely with phrase structure.

$$(31) \textit{ phrase} \rightarrow \left[\begin{array}{l} \text{DTRS} \quad \boxed{0} \textit{list}([\text{UN} \ -]) \circ \left\langle \left[\begin{array}{l} \text{UN} \quad + \\ \text{DOM} \quad \boxed{1} \end{array} \right], \dots, \left[\begin{array}{l} \text{UN} \quad + \\ \text{DOM} \quad \boxed{n} \end{array} \right] \right\rangle \\ \text{DOM} \quad \boxed{0} \circ \boxed{1} \circ \dots \circ \boxed{n} \end{array} \right]$$

We must now specify which phrases can be [UN+]. The inversion of a subject indicates that the VP is unioned. In addition, the occurrence of the higher subject between an embedded verb and its complement indicates that the downstairs VP is unioned into its mother (the VP). Thus, finite and infinitival VPs can union. Head-adjunct phrases can also union as shown by the acceptability of (23b), where the adjoined adverb is compatible with the order typical of VP[*inf*] union. On the other hand, finite sentences or head-filler phrases never union. These data follow from the following three constraints:

$$(32) \text{ a. } \textit{ phrase} \rightarrow [\text{UN} \ /-]$$

$$\text{ b. } \left[\begin{array}{l} \textit{head-comps-ph} \\ \text{HEAD} \quad \textit{extr-inv-vb} \end{array} \right] \rightarrow [\text{UN} \ +]$$

$$\text{ c. } \left[\begin{array}{l} \textit{head-adj-ph} \\ \text{HEAD} \quad \textit{extr-inv-vb} \\ \text{SUBJ} \quad \textit{nelist} \end{array} \right] \rightarrow [\text{UN} \ +]$$

(32b–c) are the only types of phrases where an *extr-inv-vb* triggers union. We use a default value in (32a) to leave open the possibility that other grammatical constructions besides ETI may trigger domain union. The constraint such that the phrases triggering union are headed by an *extr-inv-vb* accounts for the long distance inversion effect along the extraction path (13). Even when the higher verb is a subject control verb, the characteristic ordering of the inverted subject preceding an embedded complement is only possible if the embedded verb is slashed.

$$(33) \text{ a. } \begin{array}{l} \textit{celui} \quad \textit{\`a} \quad \textit{qui} \quad \textit{avait} \quad \textit{promis} \quad \textit{d' \`e} \textit{crire} \quad \textit{\`a} \quad \textit{ce} \quad \textit{sujet} \\ \textit{the-one} \quad \textit{to} \quad \textit{who} \quad \textit{had} \quad \textit{promised} \quad \textit{DE} \quad \textit{write} \quad \textit{at} \quad \textit{this} \quad \textit{topic} \\ \textbf{\textit{le directeur du labo}} \\ \textit{the boss} \quad \quad \textit{of-the} \quad \textit{lab} \\ \textit{'the person whom the boss of the lab promised to write about} \\ \textit{this topic'} \end{array}$$

$$\text{ b.*} \textit{celui} \quad \textit{\`a} \quad \textit{qui} \quad \textit{avait} \quad \textit{promis} \quad \textit{d' \`e} \textit{crire} \quad \textbf{\textit{le directeur}} \quad \textit{\`a} \quad \textit{ce} \quad \textit{sujet}$$

The analysis is illustrated in figure 4, taking the example of (13a).

LP statements. LP statements are constraints on DOM values. Two constraints are sufficient to account for the basic order pattern.

$$(34) \text{ a. } \left[\begin{array}{l} \textit{headed-phrase} \\ \text{HD-DTR} \quad \boxed{1}[\text{LEX} \ +] \end{array} \right] \rightarrow \left[\text{DOM} \quad \boxed{1} \prec [\quad] \right]$$

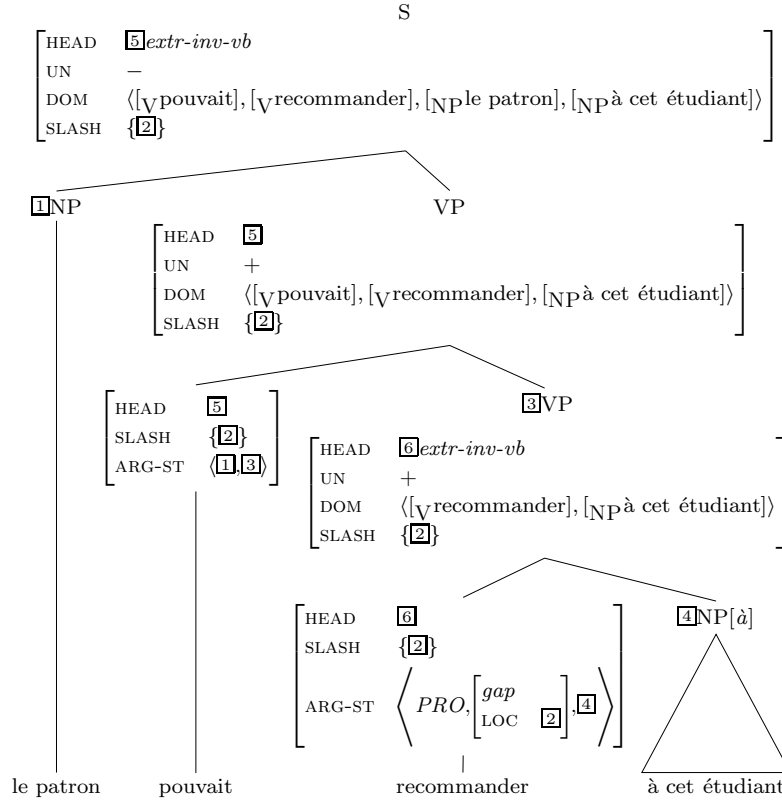


FIGURE 4 Domain union analysis of (13a)

$$b. \textit{phrase} \rightarrow \left[\text{DOM} \left[\text{HEAD} \textit{extr-inv-vb} \right] \prec \text{NP}[\textit{acc}] \right]$$

Constraint (34a) accounts for the fact that the head verb comes first in any verb-headed phrase.¹⁴ Since it applies to DOM at all phrasal levels, (34a) also ensures that the verbs of unioned constituents come in the right order: each of them comes first in the lowest domain on which it occurs, which implies that it comes before the other members of this domain on higher domains too. For instance, in figure 4, since *pouvoir* must precede *recommander* on the middle VP’s DOM, it must precede it at higher nodes in the tree too.

¹⁴This is a simplification which does not take into account the distribution of adverbs and bare quantifiers (Abeillé and Godard 1997, 1998). For a fuller presentation and account of word order constraints in ETI, see Bonami and Godard to appear.

Constraint (34b) accounts for the fact that the inverted subject must follow not only the verb selecting it, but any other inverted verb occurring on the same domain; this accounts for the contrast in grammaticality between (13b) and (35) below.

(35)*le livre que croyait **le patron du labo** pouvoir recommander à cet étudiant

1.5 Locative inversion

The analysis extends to a poorly studied class of sentences in French which look like English locative inversion, but turn out to be instances of PP topicalization.¹⁵

1.5.1 The inverted NP is a subject

The inverted NP in locative inversion (LI), beginning with a locative or temporal PP, has the very same distribution of subject-like and object-like properties as the inverted NPs in ETI. It binds *se* (36a); it does not allow quantitative *en* (36b); it does not allow bare Qs lacking an accusative form (36c); it is not compatible with floating *beaucoup* (36d); finally, it exhibits the same person/number agreement pattern (36e).

- (36) a. Dans le lac se reflétait **un château**.
 ‘In the lake was reflected a castle.’
 b.*Dans le salon en bavardaient **deux**.
 ‘In the lounge chattered two.’
 c.*Dans le salon bavardaient **quelques-uns**.
 ‘In the lounge some chattered (of men).’
 d.*Au mur étaient beaucoup accrochés **de tableaux**.
 ‘To the wall hung many pictures.’
 e. Près de l’église se trouvaient/*vous trouviez **Marie et toi**.
 ‘Near the church stood Marie and you.’

We thus conclude that if the inverted NP in ETI is an accusative subject, so is the inverted NP in LI.

1.5.2 Long-distance inversion

LI is possible in a context where a raising or even a control verb intervenes between the proposed PP and the verb selecting it. The subject

¹⁵For reasons of space, we leave aside quotative inversion whose verb can also be shown to be an *extr-inv-vb*: there is a potential unbounded dependency between the verb of ‘saying’ and the higher verb, the subject can occur between the embedded V and the embedded complement (‘*Ce n’est pas de sa faute*’, *semblait vouloir dire Marie aux policiers*. ‘“He is not responsible”, Marie seemed to want to say to the policemen’)

is then realized on the right of the embedded verb.¹⁶

- (37) a. Dans le jardin semblaient danser **des statues de pierre**.
 ‘In the garden seemed to be dancing statues of stone.’
 b. Du piano semblait vouloir s’échapper **un son métallique**.
 ‘From the piano, a metallic sound seemed to want to escape.’

As in ETI, the inverted subject of the main verb can be linearized between the constituents of an embedded VP, as (38) shows.

- (38) Sur la place semblait se dresser **une cathédrale** avec majesté.
 ‘A cathedral appeared to stand majestically on the square.’

Moreover, the properties which are problematic for a flat structure analysis also hold in LI sentences: clitics stay on the downstairs verb (39a) and the verbal complement can be ‘marked’ (39b).

- (39) a. A l’autre bout du pont semblait lui répondre **l’église de la Madeleine**.
 ‘On the other end of the bridge the church of la Madeleine seemed to match it.’
 b. Sur un côté de la place commence à se détacher **le futur opéra**.
 ‘On one side of the piazza the future opera house begins to emerge.’

We conclude that the position of the subject in ETI and in LI have the same source: in each case, the position of the subject is an effect of domain union. Since domain union can be licensed by extraction, we must examine whether LI is also an extraction construction.

1.5.3 The preposed PP is extracted

Regarding the initial PP, we are faced with two possibilities: either the PP is a filler, or the PP is not extracted; rather, it is a constituent of the VP, realized in initial position as an effect of domain union not triggered by extraction (Kathol and Levine 1993 provides such an analysis for English LI). We argue in favor of the first alternative.¹⁷

It is a fact that PP topicalization can trigger inversion (40a). This is true for locative PPs as well, independently of their argument (40b) or modifier (40c) status.¹⁸

¹⁶English sentences parallel to (37a) have been taken as evidence that the initial PP in English LI is a subject (e.g., Bresnan 1994). This argument does not hold for French, given the possible occurrence of control verbs between the PP and the verb selecting it (37b). In addition, there is no case for subject PPs in French.

¹⁷The fact that the initial PP can be semantically either an argument or a modifier is not a counter-argument against an extraction analysis. In French, the two kinds of PPs do not contrast with respect to extraction (see Hukari and Levine 1995).

¹⁸Note that extraction of PPs out of finite *wh*-clauses is possible in French (Godard

- (40) a. A ses fils, Paul pensait qu' était dévolue **la maison**.
to his sons Paul thought that was allotted the house
'Paul thought that the house was allotted to his sons.'
- b. Du château, Paul croyait que sortirait **un**
from-the castle Paul believed that come-out-COND a
groupe de soldats.
group of soldiers
'Paul believed that a group of soldiers would come out of the
castle.'
- c. Dans la cave, Paul croyait que complotait **un**
in the basement Paul believed that conspired a
groupe de terroristes.
group of terrorists
'Paul believed that a group of terrorists were conspiring in the
basement.'

Therefore an extraction analysis for LI sentences is always available in the grammar of French. As there is no recognizable ambiguity in LI sentences, there is no case for an alternative syntactic analysis.

There appear to be some interpretive constraints on subject inversion, which we leave for further research. There certainly are some cases where inversion is obligatory, but these are found with non-locative (42) as well as locative (41) PPs.

- (41) a. Dans le salon se trouvait **un groupe d' hommes**.
In the lounge SE stood a group of men
'A group of men stood in the lounge.'
- b.*Dans le salon, un groupe d'hommes se trouvait
- (42) a. Au malheur s' ajoute **la pauvreté**.
to-the misfortune SE adds the poverty
'Poverty adds further to misfortune.'
- b.*Au malheur, la pauvreté s'ajoute.

We thus conclude that PP topicalization and LI are one and the same phenomenon, and that there is no specific locative inversion construction in French: purported LI sentences are just regular cases of extraction-triggered inversion.¹⁹

1988): *Dans le salon, pourquoi se saoulait un groupe de soldats ?* 'Why was a group of soldiers getting drunk in the lounge?'

¹⁹This differs from the analysis we gave for LI in Bonami et al. 1998, of which this paper is a development and reappraisal.

1.6 Conclusion

We have argued that the inverted NP in ETI constructions is best analyzed as a subject. After showing that both transformational and argument composition analyses fail, we provide a domain union analysis. This analysis extends to all cases of extraction, including locative inversion, which is shown to be an instance of PP topicalization in French. While domain union and argument composition may appear to be competing analyses for certain data, the comparison between NP subject inversion and causative constructions suggests that both mechanisms are necessary in a grammar of French. NP subject inversion may turn out to be the only family of constructions where domain union is required in a grammar of French. However, given the frequency of use of these constructions, NP subject inversion is a major feature of French discourse.

References

- Abeillé, A. 1997. Fonction ou position objet ? (ii et fin). *Le gré des langues* 12:8–34.
- Abeillé, A. and D. Godard. 1997. The syntax of French negative adverbs. In D. Forget, P. Hirschbühler, F. Martineau, and M. Rivero, eds., *Negation and polarity*. Amsterdam: John Benjamins.
- Abeillé, A. and D. Godard. 1998. A lexical account of quantifier floating. In G. Webelhuth, J.-P. Koenig, and A. Kathol, eds., *Lexical and constructional aspects of linguistic explanation*. Stanford: CSLI Publications.
- Abeillé, A., D. Godard, and I. Sag. 1998. Two kinds of argument composition. In E. Hinrichs, A. Kathol, and T. Nakazawa, eds., *Complex predicates*. New York: Academic Press.
- Bonami, O. and D. Godard. to appear. Inversion du sujet, constituance et ordre des mots. In G. Lardreau, ed., *Cahier Jean-Claude Milner*. Paris: Verdier.
- Bonami, O., D. Godard, and J.-M. Marandin. 1998. French subject inversion in extraction contexts. In G. Bouma, G.-J. M. Kruijff, and R. T. Oehrle, eds., *Proceedings of FHCG'98*, pages 101–112. Saarbrücken.
- Bresnan, J. 1994. Locative inversion and the architecture of universal grammar. *Language* 70:72–131.
- Godard, D. 1988. *La syntaxe des relatives en français*. Paris: Ed. du CNRS.

- Hinrichs, E. and T. Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In J. Nerbonne, K. Netter, and C. Pollard, eds., *German in HPSG*, pages 11–37. Stanford: CSLI Publications.
- Hukari, T. and R. D. Levine. 1995. Adjunct extraction. *Journal of Linguistics* 31:195–226.
- Kathol, A. 1995. *Linearization-based German syntax*. Ph.D. thesis, Ohio State University.
- Kathol, A. and R. D. Levine. 1993. Inversion as a linearization effect. In *Proceedings of NELS 23*, pages 207–221.
- Kayne, R. S. 1972. Subject inversion in French interrogatives. In J. Casagrande and B. Saciuk, eds., *Generative studies in Romance languages*, pages 70–126. Rowley: Newbury House.
- Kayne, R. S. and J.-Y. Pollock. 1978. Stylistic inversion, successive cyclicity and move NP in French. *Linguistic Inquiry* 9:595–621.
- Lambrecht, K. and M. Polinsky. 1997. Typological variation in sentence-focus constructions. In *CLS 33*.
- Le Bidois, R. 1950. *L'inversion du sujet dans la prose contemporaine*. Paris: Editions d'Artey.
- Marandin, J.-M. 1997. Dans le titre se trouve le sujet. Mémoire d'habilitation, Université Paris 7.
- Miller, P. 1992. *Clitics and constituents in phrase structure grammar*. New York: Garland.
- Milner, J.-C. 1978. Cyclicité successive, comparatives et cross over en français. *Linguistic Inquiry* 9:673–693.
- Reape, M. 1994. Domain union and word-order variation in German. In J. Nerbonne, K. Netter, and C. Pollard, eds., *German in HPSG*, pages 151–197. Stanford: CSLI Publications.
- Sag, I. A. 1997. English relative clause constructions. *Journal of Linguistics* 33:431–484.
- de Wind, M. 1995. *Inversion in French*. Ph.D. thesis, Groningen University.
- Zaenen, A. 1983. On syntactic binding. *Linguistic Inquiry* 14:469–504.