

# Topic vs. Comment in some Subject Inversion Sentences in French and Portuguese

MANUELA AMBAR  
JEAN-YVES POLLOCK

## Abstract

*In Portuguese, though not in French, postverbal subjects provide new information, i.e. in the unmarked case postverbal subjects are foci in Portuguese, not topics, as claimed in Ambar (1988). However, in Portuguese or French wh-questions the postverbal subject does not provide or ask for new information. These postverbal subjects are not foci. These facts lead one to raise the following two questions: (i) why is it that Portuguese syntax allows for a postverbal topic in wh-questions, an option it typically bans elsewhere?; (ii) why does French only allow for postverbal subjects in stylistic inversion sentences where the postverbal subject is a topic? We shall attempt to provide a unified answer to these two questions here. In line with Kayne & Pollock (1998)-(2001), our main claim will be that despite appearances a topic DP does stand in the left periphery of the input structures in the two languages. That topic position in the CP domain is analogous, though not identical, to the position in which a clitic left dislocated DP stands. The postverbal occurrence of the (topic) subject must result from further remnant movement of the whole IP to a position past the topicalized subject.*

## 1. Introduction

In neither (1) nor (2),

- (1) A quem falou o João?  
'To whom spoke João?'
- (2) À qui a parlé Jean?  
'To whom spoke Jean?'

does the postverbal subject provide or ask for new information: *o João* and *Jean* are not *foci*; in such sentences only the *wh-* words *a quem* and *à qui* are *foci*. However, in contexts like (3) Portuguese, though not French, as the ungrammaticality of (4) shows, does allow for a postverbal focus subject:

(3) Falou o João

(4) \*A parlé Jean

here *o João* does provide new information since (3) is a fine answer to the question *Quem falou?* ‘Who talked?’. The ill-formedness of (4) in Modern French should thus probably be tied to the impossibility of focusing postverbal subjects in environments different from (5),

(5) N’a parlé que Jean  
neg has spoken that Jean  
‘Only Jean spoke’

where the DP in *que+DP* strings is interpreted as the only licit element from among a set of invalid alternatives: in (5) *Jean* is a *contrastive* focus.

Elementary facts such as these may lead one to raise the following two questions:

**A.** Why is it that Portuguese syntax allows for a postverbal topic in (1), an option it typically bans elsewhere, as (3) shows?

**B.** Why does French only allow for postverbal subjects in stylistic inversion sentences where the postverbal subject is a topic?

We shall attempt to provide a unified answer to these two questions here. In line with Kayne & Pollock (1998)-(2001), our main claim will be that despite appearances a topic DP does stand in the left periphery of the input structures to (1) and (2) in the two languages. That topic position in the CP domain is analogous, though not identical, to the position in which the clitic left dislocated DP of sentences like (6) in French or Italian stand:

- (6) a. Pierre, il a parlé  
Pierre, he spoke  
b. Pierre, je l’ai vu  
Pierre, I him saw  
c. Piero, l’ho visto  
Piero, (I) him saw

If so, the postverbal occurrence of the (topic) subject must result from further *remnant movement* of the whole IP to a position past the topicalized subject, as the derivation in (7) sketches:<sup>82</sup>

- (7) a. Pierre a parlé quand  
 Topicalization of subject DP  $\Rightarrow$   
 b. [<sub>TopP</sub> Pierre<sub>i</sub> Top° [<sub>IP</sub> t<sub>i</sub> a parlé quand]]  
 Remnant IP movement  $\Rightarrow$   
 c. [<sub>GP</sub> [<sub>IP</sub> t<sub>i</sub> a parlé quand]<sub>j</sub> G° [<sub>TopP</sub> Pierre<sub>i</sub> Top° t<sub>j</sub> ]]  
 Wh- Movement  $\Rightarrow$   
 d. [<sub>WhP</sub> Quand<sub>k</sub> Wh° [<sub>GP</sub> [<sub>IP</sub> t<sub>i</sub> a parlé t<sub>k</sub> ]<sub>j</sub> G° [<sub>TopP</sub> Pierre<sub>i</sub> Top° t<sub>j</sub> ]]

In the present work we shall provide empirical evidence in favour of step (7b) and attempt to shed light on the relationship between (7c) and (7d). In short, we shall try to say why Wh- Movement makes Remnant IP movement past the topic position licit in the two languages, thereby providing the first step to a unified answer to questions A and B.

## 2. Some Arguments for Remnant IP Movement in French wh-Questions

In the framework sketched in (7) for French Stylistic Inversion no IP constituent can c-command out of IP in structure (8):

- (8) [<sub>GP</sub> IP<sub>i</sub> G° [ DP t<sub>i</sub> ]]

This general structural property can be made use of to account for the fact that floating quantifiers cannot be extracted from postverbal subjects of stylistic inversion sentences, as the ungrammaticality of (9a, c, e, g) shows.<sup>83</sup>

- (9) a. \*Le jour où ont {peu, beaucoup, trop} téléphoné de linguistes  
 The day when have {few, many, too many} telephoned of linguists  
 ‘The day when {few, many, too many} linguists telephoned’  
 b. Le jour où ont téléphoné {peu, beaucoup, trop} de linguistes  
 The day when have telephoned {few, many, too many} of linguists  
 ‘The day when {few, many, too many} linguists telephoned’  
 c. \*Quand se sont {peu, beaucoup, trop} trompés de linguistes?  
 When refl. are {few, many, too many} mistaken of linguistes?  
 ‘When have {few, many, too many} linguists made mistakes?’  
 d. Quand se sont trompés {peu, beaucoup, trop} de linguistes?  
 When refl. are mistaken {few, many, too many} of linguistes?  
 ‘When have {few, many, too many} linguists made mistakes?’

<sup>82</sup>On ‘GP’ see e.g. Kayne & Pollock (2001).

<sup>83</sup>On this see also Obenauer (1984), Kayne (1984, chapter 3, appendix).

- e. ?\*Où sont {peu, beaucoup, trop} partis de linguistes?  
Where are {few, many, too many} gone of linguists?  
'Where have {few, many, too many} linguists gone?'
- f. Où sont partis {peu, beaucoup, trop} de linguistes?  
Where are gone {few, many, too many} of linguists?  
'Where have {few, many, too many} linguists gone?'
- g. ?\*Dans quel congrès ont {peu, beaucoup, trop} été récompensés de linguistes?  
In what conference have {few, many, too many} been rewarded of linguists?  
'In what conference have {few, many, too many} linguists been rewarded?'
- h. Dans quel congrès ont été récompensés {peu, beaucoup, trop} de linguistes?  
In what conference have been congratulated {few, many, too many} of linguists?  
'In what conference have {few, many, too many} linguists been rewarded?'

Although Obenauer's (1984) 'Quantification at a Distance' – henceforth QAD – can licitly extract quantifiers like *peu*, *beaucoup*, *trop* from a direct object or from the postverbal subject of *il* impersonal constructions,<sup>84</sup> it cannot do so from the postverbal subject of SI sentences, regardless of the type of verb used in the sentence.<sup>85</sup> In short, postverbal subjects behave like preverbal

---

<sup>84</sup>As in (i) and (ii) or (16) in the text below:

- (i) On a {trop, tant, beaucoup} récompensé de linguistes  
one has {too many, so many, many} congratulated of linguists  
'People have congratulated {so many, too many, many} linguists'
- (ii) Il a {trop, tant, beaucoup} été récompensé de linguistes  
it has {too many, so many, many} been congratulated of linguists  
'There have been {too many, so many, many} linguists congratulated'

<sup>85</sup>As (9) shows, there are variations in inacceptability in (9) and the like, for some inaccusatives like *partir*, pronominals like *se tromper* and passive participles like *récompensés* yield somewhat less unacceptable cases of QAD than inergatives like *téléphoner*. In the spirit of Kayne (1984, chapter 3 note 61), one could suggest that these variations arise when SI configurations do not result from Remnant IP movement at all but are parasitic on *il* impersonal constructions (see previous footnote and text below). It could be claimed for example that for such speakers (9e, g) have a non lexical *il* expletive in subject position and a postverbal associate in an object-like position. This does not jibe well with Kayne & Pollock (2001, (text to) note 31), however, who suggest rather that all such examples do indeed involve remnant IP movement and that French never allows for null expletives. If so, one might want to extend to QAD with inaccusatives, pronominal verbs and passive participles what they suggest (see their note 9) for *en* extraction – taken up in a slightly different form in the text below (see also next footnote) – and say that some speakers marginally allow QAD to take place from a postverbal position in IP before

subjects with respect to QAD: neither allows for movement into a non c-commanding position.

It is worth pointing out that postverbal subjects do allow for subextraction in other cases, like (10) for example, where the wh- quantifier *combien* has stranded *de linguistes*:

- (10) a. *Combien ont téléphoné de linguistes?*  
How many have phoned of linguists?  
'How many linguists have phoned?'
- b. *Combien sont partis de linguistes?*  
How many are gone of linguists?  
'How many linguists have gone?'
- c. *Combien se sont trompé de linguistes?*  
How many refl. are mistaken of linguists?  
'How many linguists made mistakes?'
- d. *Combien ont été récompensé de linguistes?*  
How many were rewarded of linguists?  
'How many linguists were rewarded?'

Since movement must always be to a c-commanding position, pairs like (9) vs. (10) follow from the Remnant IP movement analysis of SI: while *peu*, *beaucoup* and *trop* fail to c-command their trace in the postverbal subject from their IP internal position, the wh- phrase *combien* does, since it has moved 'higher up' in the left periphery, say to the specifier position of WhP as illustrated in (11) and (12):

- (11) \*[ ... [GP [IP t<sub>i</sub> ont [QP peu]<sub>i</sub> téléphoné]<sub>j</sub> G<sup>o</sup>[TopP [DP t<sub>i</sub> de linguistes] Top<sup>o</sup> t<sub>j</sub> ]
- (12) [WhP combien<sub>i</sub> wh<sup>o</sup> [GP [IP t<sub>i</sub> ont téléphoné]<sub>j</sub> G<sup>o</sup>[TopP [DP t<sub>i</sub> de linguistes] Top<sup>o</sup> t<sub>j</sub> ]

This analysis carries over to (14) and (13) on Kayne's (1984, chapter 3) well supported assumption that in (13a) and the like *de NP* contains a null QP that *pas* must c-command at spell-out.

- (13) a. *Je n'ai pas mangé de pain*  
I neg have not eaten of bread  
'I haven't eaten any bread'
- b. \**De pain n'a pas été mangé*  
of bread neg has not been eaten  
'No bread has been eaten'

---

remnant IP movement does; QAD would then not be movement to a non c-commanding position, whence the less degraded quality of (9e, g).

- (14) a. \**Quel livre n'ont pas lu de linguistes?*  
 What book neg have not read of linguists?  
 'What book has no linguist read?'
- b. \**Le livre que ne comprennent pas de linguistes, c'est le mien*  
 The book that neg understand not of linguists, it's mine  
 'My book is the book that no linguist understands'

Examples like (15), which are very close synonyms of (14), are perfect. This is because a DP like *aucun linguiste* does not contain an empty category that a displaced quantifier must bind.

- (15) a. *Quel livre ne comprend aucun linguiste?*  
 What book neg understands no linguist?  
 'What book does no linguist understand?'
- b. *Le livre que ne comprend aucun linguiste, c'est le mien*  
 The book that neg understand no linguist, it's mine  
 'My book is the book that no linguist understands'

As we have already pointed out, *il* impersonal constructions do allow for subextraction from the expletive's associate as well as binding by *pas* of the associate's empty QP:

- (16) a. *Il n'est pas entré de linguistes dans cette pièce depuis ce matin*  
 it neg is not come in of linguists in this room since this morning  
 'No linguist has come into that room since this morning'
- b. *Il n'a pas téléphoné de linguiste pour toi ce matin*  
 it neg has not telephoned of linguist for you this morning  
 'No linguist has rung you up this morning'
- c. *Il est {peu, beaucoup, trop} entré de linguistes dans cette pièce depuis ce matin*  
 it is not {few, (too) many} come in of linguists in this room since this morning  
 '{few, (too) many} linguists have come into that room since this morning'
- d. *Il a {peu, beaucoup, trop} téléphoné de linguistes pour toi ce matin*  
 it has {few, (too) many} telephoned of linguist for you this morning  
 '{few, (too) many} linguists have rung you up this morning'

- e. Il (ne) s'est {pas, peu, beaucoup, trop} trompé de linguistes ce matin  
 it (neg) refl is {not, few, (too) many} mistaken of linguist this morning  
 ' {No, few, (too) many} linguist(s) has/have been mistaken this morning'
- f. Il (ne) est {pas, peu, beaucoup, trop} parti de linguistes ce matin  
 it (neg) is {not, few, (too) many} gone of linguist this morning  
 ' {No, few, many, too many} linguist(s) has/have gone this morning'
- g. Il (n') a {pas, peu, beaucoup, trop} été récompensé de linguistes  
 it (neg) has {not, few, (too) many} been rewarded of linguist(s)  
 ' {No, few, (too) many} linguist(s) has/have been rewarded'

In such sentences, no Remnant IP movement has taken place; so nothing blocks extraction from the postverbal associate of *il*, or binding of one of its (null) constituents.

So called 'quantitative *en*' – on which see Pollock (1986, 1998), Boivin (2000) – can cliticize freely from a direct object, as in (17), as well as from the postverbal subject of *il* impersonal constructions, as in (18):

- (17) Le jour où le juge en a condamné {trois, peu, beaucoup, trop}  
 The day where the judge of them has condemned {three, few, (too) many}  
 'The day when the judge condemned {three, few, (too) many} of them'
- (18) a. Le jour où il en est parti {trois, peu, beaucoup, trop}  
 The day where it of them is gone {three, few, many, too many}  
 'The day when {three, few, (too) many} of them left'
- b. Dans quelle congrès en a-t-il été récompensés {trois, peu, beaucoup, trop}?  
 In what conference of them have been rewarded {three, few, (too) many}?  
 'In what conference {three, few, (too) many} of them were rewarded?'

Yet no quantitative *en* can be extracted from a preverbal subject – (19) – or from the postverbal subject of Stylistic Inversion sentences – (20):

- (19) a. \*{trois, peu, beaucoup, trop} en ont téléphoné  
 {three, few, many, too many} of them+have phoned  
 ' {three, few, many, too many} of them have phoned'
- b. \*Je ne sais pas à qui {trois, peu, beaucoup, trop} en+ont téléphoné

- I neg know not to whom {three, few, (too) many} of them have phoned  
 'I don't know to whom {three, few, (too) many} of them have phoned'
- c. \*{trois, peu, beaucoup, trop} en sont partis  
 {three, few, many, too many} of them+are gone  
 '{three, few, (too) many} of them have gone'
- d. \*Je ne sais pas dans quelle ville {trois, peu, beaucoup, trop} en ont disparu  
 I neg know not in what city of them+are disappeared {three, few, many, too many}  
 'I don't know in what city {three, few, (too) many} of them disappeared'
- (20) a. \*Le jour où en ont téléphoné {trois, peu, beaucoup, trop}  
 The day where of them+have phoned {three, few, (too) many}  
 'The day when {three, few, (too) many} of them phoned'
- b. \*Je ne sais pas à qui en+ont téléphoné {trois, peu, beaucoup, trop}  
 I neg know not to whom of them have phoned {three, few, (too) many}  
 'I don't know to whom {three, few, many, too many} of them phoned'
- c. \*Le jour où en sont partis {trois, peu, beaucoup, trop}  
 The day whereof them+are gone {three, few, (too) many}  
 'The day when {three, few, many, too many} of them left'
- d. \*Je ne sais pas dans quelle ville en ont disparu {trois, peu, beaucoup, trop}  
 I neg know not in what city of them+are disappeared {three, few, (too) many}  
 'I don't know in what city {three, few, (too) many} of them disappeared'

The only well-formed counterparts of (19) would contain an elliptical subject, as in (21),

- (21) a. {trois, peu, beaucoup, trop} ont téléphoné  
 {three, few, many, too many} have phoned
- b. Je ne sais pas à qui {trois, peu, beaucoup, trop} ont téléphoné  
 I neg know not to whom {three, few, (too) many} have phoned  
 'I don't know to whom {three, few, (too) many} of them have phoned'
- c. {trois, peu, beaucoup, trop} sont partis  
 {three, few, (too) many} are gone  
 '{three, few, (too) many} have gone'

- d. Je ne sais pas dans quelle ville {trois, peu, beaucoup, trop} ont disparu  
 I neg know not in what city of them+are disappeared {three, few, (too) many}  
 'I don't know in what city {three, few, (too) many} have disappeared'

but such elliptical subjects cannot occur as the postverbal subjects of SI structures:

- (22) a. \*La femme à qui ont téléphoné {trois, peu, beaucoup, trop}  
 The woman to whom have phoned {three, few, (too) many}  
 'The woman to whom {three, few, (too) many} have phoned'  
 b. \*Le livre qu'ont lu {trois, peu, beaucoup, trop}  
 The book that have read {three, few, (too) many}  
 'The book that {three, few, (too) many} have read'  
 c. \*Je ne sais pas dans quelle ville se sont installés {trois, peu, beaucoup, trop}  
 I neg know not in what city refl are settled {three, few, (too) many}  
 'I don't know in what city {three, few, (too) many} have settled'

Let us first try to account for the ban on quantitative *en* cliticization from the postverbal subject of stylistic inversion sentences. Our analysis so far evidently invites us to look at (20) in the same light as the ungrammaticality of QAD cases like (9); on that view, no *en* cliticization would be allowed to take place from the topicalized subject of SI structures if Remnant IP movement takes place before *en* cliticization does, since that would again be movement to a non c-commanding position, just as in QAD cases.

Taking that tack would however leave us without an account for why so-called 'adnominal' *en* CAN be extracted both from canonical preverbal subjects and postverbal subjects in SI, as (23) shows:

- (23) a. Le premier chapitre de ce livre {sera publié, paraîtra} dans cette revue  
 The first chapter of this book {will be published, will appear} in that journal  
 b. Le premier chapitre en {sera publié, paraîtra} dans cette revue  
 The first chapter of it+{will be published, will appear} in that journal

- c. Dans quelle revue *en* {sera publié, paraîtra} le premier chapitre?  
 In what journal of it+{will be published, will appear} the first chapter?

As in Pollock (1998), we shall claim that the main difference between the two types of *en* lies in the fact that the DP from which *quantitative en* is extracted is elliptical in a way that the DP out of which *adnominal en* moves isn't. More precisely, we hold that the relevant input structures to *quantitative en* and *adnominal en* are (24a) and (24b) respectively:

- (24) a. [QP {trois, peu, beaucoup} [NP [pro] [PP *en*]]  
 b. [DP *le* [NP [AP premier] [N chapitre] [PP *en*]]]

Let us say further, again as in Pollock (1998), that the content of *pro* in (24a) must be 'recovered' at PF, a requirement we interpret as meaning that *pro* must have formal features associated with it. If it failed to do so, *pro* would be 'unreadable' by the PF component, which would cause the derivation to crash; we shall say that *pro* can 'inherit' the relevant features by being anaphoric to an item with the relevant features. Suppose *en* is the only possible antecedent for *pro* in *quantitative en* cases. (20) now follows since in such examples *en* is NOT standing in the c-commanding position that would allow it to bind *pro* and provide it with the identifying formal features it needs. The same will be true of (20), exactly for the same reason under the remnant movement analysis of SI. In *quantitative en* cases like (23), however, no *pro* is present, so the final position of *en* is immaterial to the PF licensing of a fully specified DP like *le premier chapitre*.<sup>86</sup>

We can now go back to pairs like (22) vs. (23); everything else being equal, we must assume that in (22) the elliptical subject must have been 'licensed', i.e. must have had formal features associated with it, which in turn requires that a

<sup>86</sup>We are thus claiming that, in both *adnominal* and *quantitative en* cliticization, *en* moves to its clitic position in IP *before* IP is remnant moved, as the derivations in (i) and (ii) sketch:

- (i) Input: [V sera [SC [le premier chapitre *en*] publié]]  
*en* cliticization ⇒  
 a [IP *en*<sub>i</sub> [V sera [SC [DP *le premier chapitre* t<sub>i</sub>] publié]]  
 [...] and DP to Spec IP (EPP) ⇒  
 b [IP [DP *le premier chapitre* t<sub>i</sub>] <sub>j</sub> *en*<sub>i</sub> [V sera [SC t<sub>j</sub> publié]]
- (ii) Input: [V sera [SC [DP un [pro *en*]] publié]]  
*en* cliticization ⇒  
 a [IP *en*<sub>i</sub> [V sera [SC [DP un pro t<sub>i</sub>] publié]]  
 [...] and DP to Spec IP (EPP) ⇒  
 b [IP [DP un pro t<sub>i</sub>] *en*<sub>i</sub> [V sera [SC publié]]

What goes wrong in (iib) is that *pro* cannot recover its formal features at PF from non c-commanding *en*, as it must.

local c-commanding element bind it at PF. We shall posit that the element in question is a D(iscourse)-linked (null) Topic operator standing in one of the outer layers of the Split CP (cf. e.g. Ambar (1988), Rizzi (1997), Poletto (1998)), as sketched in (25) where the shared index informally notates anaphoric binding of *pro* by the D-linked topic Operator.

(25) [<sub>TopP</sub> OP<sub>i</sub> Top° [<sub>IP</sub> [<sub>QP</sub> beaucoup [<sub>NP</sub> pro]<sub>i</sub> ] ont téléphoné]]

On Kayne and Pollock's Remnant movement analysis of SI, on the other hand, the relevant subpart of the spell-out parse of illicit SI cases like (23) would be as in (26):

(26) [<sub>TopP</sub> OP Top° [<sub>GP</sub> [<sub>IP</sub> t<sub>i</sub> ont téléphoné]<sub>j</sub> G° [<sub>TopP</sub> [<sub>QP</sub> [ beaucoup] [<sub>pro</sub>]<sub>i</sub> t<sub>j</sub> ]]]

We claim that the Remnant IP that has moved to the specifier of GP in SI sentences is 'blocking' local binding of *pro* by the null Topic Operator in the leftmost Topic phrase: OP and *pro* are 'too far apart', separated as they are by a (remnant) moved IP. If so, the elliptical QP cannot inherit the features it needs for lexicalisation and the derivation crashes because it ends up containing an uninterpretable item at PF.<sup>87</sup>

In sum, in this section we have claimed with Kayne & Pollock (1988), (2001) that the island properties of postverbal subjects in SI contexts are best understood as resulting from improper movement into a remnant moved IP standing in a 'higher' layer of the split CP of French and Romance. This allows for a neat account of the ban of QAD from postverbal subjects in SI and of the minimally different extraction of *combien* from SI postverbal subjects; independent properties of *en* cliticization combine with this account of SI to yield the complex array of judgments that speakers pass on adnominal and quantitative *en* extraction from preverbal and postverbal subjects.

---

<sup>87</sup> As above we say that the PF structure of (26) contains an object that cannot lexicalise because it does not have the (formal) features that this would require; put another way, even though *pro* doesn't have a phonetic matrix, it must still obey constraints on lexicalisation, which requires at least formal features. If we are to account for pairs like (i) vs. (ii),

- (i) \*Trois en sont venus
- (ii) Trois sont venus

it has to be the case that in a NP merged as [<sub>NP</sub> pro [<sub>PP</sub> en]] *pro* cannot have its feature content licensed by a c-commanding topic operator. In sum, in elliptical quantitative DPs French allows for either identification of *pro* by *en* or by a null topic operator but the two strategies cannot be used at the same time, for (economy) reasons that still need to be fully understood.

### 3. Portuguese

Arguments for Remnant IP movement in Portuguese questions based on this type of movement cannot be constructed in Portuguese, unfortunately, since Portuguese does not have any QAD constructions or binding of a null QP by some counterpart of *pas* or an adnominal-quantitative clitic like *en*. However, it is possible to use another paradigm to push the same idea: non specific subject indefinites are sharply excluded from sentence final positions. The following contrasts illustrate the phenomenon we have in mind:

- (27) a. *Em que festa o Pedro comeu caril?*  
At what party the Pedro ate curry?  
'At what party did Pedro eat curry?'
- b. *Em que festa alguém comeu caril?*  
At what party someone ate curry?  
'At what party did someone eat curry?'
- (28) a. *Em que festa comeu caril o Pedro?*  
At what party ate curry the Pedro?  
'At what party did Pedro eat curry?'
- b. *\*Em que festa comeu caril alguém?*  
At what party ate curry someone?  
'At what party did someone eat curry?'
- (29) a. *Em que festa comeu o Pedro caril?*  
At what party ate the Pedro curry?  
'At what party did Pedro eat curry?'
- b. *Em que festa comeu alguém caril?*  
At what party ate someone curry?  
'At what party did someone eat curry?'
- (30) a. *Em que festa tinha o Pedro comido caril?*  
At what party has the Pedro eaten curry?  
'At what party has Pedro eaten curry?'
- b. *Em que festa tinha alguém comido caril?*  
At what party has someone eaten curry?  
'At what party has someone eaten curry?'
- (31) a. *Que cartaz colou na parede o Pedro?*  
What poster pasted on the wall the Pedro?  
'What poster did Pedro paste on the wall?'
- b. *\*Que cartaz colou na parede alguém?*  
What poster pasted on the wall someone?  
'What poster did someone paste on the wall?'

- (32) a. Que cartaz colou o Pedro na parede?  
What poster pasted the Pedro on the wall?  
'What poster did Pedro paste on the wall?'
- b. Que cartaz colou alguém na parede?  
What poster pasted the Pedro on the wall?  
'What poster did Pedro paste on the wall?'
- (33) a. Que cartaz tinha o Pedro colado na parede?  
What poster has the Pedro pasted on the wall?  
'What poster has Pedro pasted on the wall?'
- b. Que cartaz tinha alguém colado na parede?  
What poster has someone pasted on the wall?  
'What poster has someone pasted on the wall?'

We believe that those facts should be seen in the same light as contrasts like (34) in French, first discussed in Cornulier (1974):

- (34) a. \*Quelle affiche a collé quelqu'un?  
What poster has pasted someone?  
'What poster did someone paste?'
- b. Quelle affiche quelqu'un a-t-il collée?  
What poster someone has he pasted?  
'What poster has someone pasted?'

Such facts show that an intrinsically indefinite subject DP like *quelqu'un* or *alguém* cannot occur in sentence final position. This rather surprising fact can be neatly accounted for on our analysis: assuming SI does involve Remnant IP movement *quelqu'un* must first topicalize to the left periphery. But it is well-known that indefinites cannot be topics, as the unacceptable (35) show:

- (35) a. \*Quelqu'un, il a téléphoné  
Someone, he has phoned
- b. \*Alguém, vi-o no cinema  
Someone, I have seen him at the cinema

On the other hand, it is also well-known that in 'complex inversion' construction like (34b) the preverbal DP is standing in an ordinary subject-like position which allows for all indefinites, as (36) and the like show.

- (36) a. Rien n'est-il certain?  
Nothing neg is it certain?  
'Isn't anything certain?'

- b. *Personne n'a-t-il compris?*  
 No one neg has-he understood?  
 'Hasn't anyone understood?'

The Portuguese facts in (27) through (33) can be analysed in a completely parallel fashion if sentences like (28) and (31) are also derived via Remnant IP movement, as sketched in (37):

- (37) Input: *o Pedro comeu caril em que festa?*  
 Subject DP Topicalization  $\Rightarrow$
- $[\text{TopP } o \text{ Pedro}_i \text{ Top}^\circ [\text{IP } t_i \text{ comeu caril em que festa}]]$   
 Remnant IP movement to GP  $\Rightarrow$
  - $[\text{GP } [\text{IP } t_i \text{ comeu caril em que festa}]_j \text{ G}^\circ[\text{TopP } o \text{ Pedro}_i \text{ Top}^\circ t_j ]]$   
 Wh movement  $\Rightarrow$
  - $[\text{WhP } [\text{em que festa}]_k \text{ Wh}^\circ [\text{GP } [\text{IP } t_i \text{ comeu caril } t_k ]_j \text{ G}^\circ[\text{TopP } o \text{ Pedro}_i \text{ Top}^\circ t_j ]]]$
- (38) Input: *alguém comeu caril em que festa?*  
 Subject DP Topicalization  $\Rightarrow$
- $[\text{TopP } algu\acute{e}m_i \text{ Top}^\circ [\text{IP } t_i \text{ comeu caril em que festa}]]$   
 Remnant IP movement to GP  $\Rightarrow$
  - $[\text{GP } [\text{IP } t_i \text{ comeu caril em que festa}]_j \text{ G}^\circ[\text{TopP } algu\acute{e}m_i \text{ Top}^\circ t_j ]]$   
 Wh movement  $\Rightarrow$
  - $[\text{WhP } [\text{em que festa}]_k \text{ Wh}^\circ [\text{GP } [\text{IP } t_i \text{ comeu caril } t_k ]_j \text{ G}^\circ[\text{TopP } algu\acute{e}m_i \text{ Top}^\circ t_j ]]]$

The first step in the derivation in (38) yields an impossible structure, since the indefinite *alguém* cannot be attracted to TopP. As for (29)-(30), we shall posit that they are the Portuguese counterparts to French Subject Clitic Inversion or Complex Inversion; it is generally thought that such V2 configurations arise as a result of (V to) Infl to C°, as indicated in (39):

- (39) Input:  $[\text{InflP } il [\text{Infl}^\circ a_i] [ t_i [ \text{parlé quand } ]]]$   
 Infl° to C°  $\Rightarrow$
- $[\text{CP } [\text{Infl}^\circ a_i]_j + \text{C}^\circ [\text{InflP } il t_j [ t_i [ \text{parlé quand } ]]]]$   
 Wh movement  $\Rightarrow$
  - $[\text{WhP } quand_k \text{ wh}^\circ [\text{CP } [\text{Infl}^\circ a_i]_j + \text{C}^\circ [\text{InflP } il t_j [ t_i [ \text{parlé } t_k ]]]]]]$

We extend this analysis to Portuguese as shown in (40):

- (40) Input:  $[\text{InflP } o \text{ Pedro } [\text{Infl}^\circ tinha_i] [ t_i [ \text{colado que cartaz na parede } ]]]$   
 Infl° to C°  $\Rightarrow$

- a. [CP [Infl<sup>o</sup> tinha<sub>i</sub>]<sub>j</sub>+C<sup>o</sup> [InflP o Pedro t<sub>j</sub> [ t<sub>i</sub> [colado que cartaz na parede ]]]]  
Wh movement ⇒
- b. [WhP [que cartaz]<sub>k</sub> wh<sup>o</sup>[CP [Infl<sup>o</sup> tinha<sub>i</sub>]<sub>j</sub>+C<sup>o</sup> [InflP o Pedro t<sub>j</sub> [ t<sub>i</sub> [ colado t<sub>k</sub> na parede]]]]]

In neither (39) nor (40) does the subject move to a topic position, therefore the indefinite *alguém* can substitute for *o Pedro* in (40), yielding the well-formed (33b), derived as in (41):

- (41) Input: [InflP alguém [Infl<sup>o</sup> tinha<sub>i</sub>] [ t<sub>i</sub> [ colado que cartaz na parede ]]]  
Infl<sup>o</sup> to C<sup>o</sup> ⇒
- a. [CP [Infl<sup>o</sup> tinha<sub>i</sub>]<sub>j</sub>+C<sup>o</sup> [InflP alguém t<sub>j</sub> [ t<sub>i</sub> [colado que cartaz na parede ]]]]  
Wh movement ⇒
- b. [WhP [que cartaz]<sub>k</sub> wh<sup>o</sup>[CP [Infl<sup>o</sup> tinha<sub>i</sub>]<sub>j</sub>+C<sup>o</sup> [InflP alguém t<sub>j</sub> [ t<sub>i</sub> [ colado t<sub>k</sub> na parede]]]]]

#### 4. Some Ambiguous Structures

The preceding section has distinguished two types of wh- questions. The first type involves topicalization of the subject DP and Remnant IP movement to the left periphery; it bans any attraction to IP of an element from within the topicalized subject, as this would be movement to a non c-commanding position, and it accounts for the ‘anti-indefiniteness’ effect that French and Portuguese sentences like (34a) and (28-31b) show. The second type, made familiar by much past work on V2 languages and constructions, involves head to head movement of I<sup>o</sup> to C<sup>o</sup> and does not involve any movement of the subject DP to a topic position in the left periphery, which can therefore be an indefinite.<sup>88</sup> The two constructions are further illustrated in (42) in the two languages, where (42c,d) are examples of the so-called ‘complex inversion’ specific to French:

- (42) a. Em que serviço tinha alguém entregado os documentos?  
To what department has someone brought back the documents?
- b. A que director tinha alguém lido o relatório?  
To what head of department has someone read the report?
- c. À quel service quelqu’un a-t-il rapporté les papiers?  
(same as (42a))
- d. A quel chef de service quelqu’un a-t-il lu le rapport?  
(same as (42b))

<sup>88</sup>On an alternative to the head movement analysis of SCLI and CI adopted here see Pollock (2000).

In addition, this framework provides an answer to question A above, repeated below:

**A.** Why is it that Portuguese syntax allows for a postverbal topic in questions like (43), an option it typically bans elsewhere?

- (43) *A quem tinha entregado os documentos o João?*  
To whom has brought back the documents the João?

What we have been saying is that the postverbal *o João* can only be a postverbal topic because the rest of the IP has moved across it on its way to the left periphery; put another way, the surprising sentence final occurrence of the topic subject in (43) and the like is misleading: here too it is standing in the position in the left periphery typically devoted to topic interpretation, despite appearances. We therefore claim that in (44) where *o João* is also a topic Remnant IP has also moved to the left periphery even though an  $I^\circ$  to  $C^\circ$  is also *a priori* possible:

- (44) *A quem falou o João?*  
To whom spoke the João?  
'To whom did João speak?'

Indeed, once the peculiar pragmatics of wh- questions is taken into account, a sentence like (45) is fine, which will follow if (45) is derived via  $I^\circ$  to  $C^\circ$ .

- (45) *A quem se queixou alguém?*  
To whom complained someone?  
'To whom did someone complain?'

The idea that strings like (44) and (45) are structurally ambiguous is supported by another paradigm involving pronouns (cf. Ambar (1988, 59, (18) and *passim*).

- (46) a. *\*?Onde tinha posto os quadros ele?*  
Where has put the paintings him?  
'Where has he put the paintings?'  
b. *A quem se queixou ele?*  
To whom complained he?  
'To whom did he complain?'

We believe that the ungrammaticality of (46a) is to be analysed as that of (47) in French,

- (47) \*Quand est parti il?  
When is gone he?

which is itself to be tied to the well-known fact that weak or clitic pronouns cannot occur in a (CLLD) topic position:

- (48) a. \*Il, il est parti  
He, he's gone  
b. ?\*Ele, ele partiu<sup>89</sup>  
(same)

Assuming so, (46a) immediately follows since it would require *ele* to move to a topic position of the same ilk as (48b); in (46b), on the other hand, *ele* is merely standing in the usual specifier of IP position and the usual I° to C° derivation can proceed smoothly, just as it can in the French analogue in (49):

- (49) De quoi se plaint-il?  
Of what refl complains he?  
'What is he complaining about?'

In sum, our account of the syntax of indefinites in Wh- questions finds independent support in the distribution of weak/clitic pronouns.

## 5. Concluding Remarks

The analysis put forth in this article finally provides us with the first ingredients of an answer to question B, repeated below:

**B.** Why does French only allow for postverbal subjects in stylistic inversion sentences where the postverbal subject is a topic?

What we have been saying in effect is that in both Portuguese and French the topic position is preverbal, more precisely an outer layer of the left periphery (of questions), and that SI sentences are only apparent counterexamples to this generalization because IP has moved across that topic position to a 'higher' ('GP') layer.

Just as in Portuguese in the cases examined in this article, Remnant IP movement to the left periphery in French is dependant on the presence of a wh-phrase in IP. This correlation can be neatly derived if one assumes that in

---

<sup>89</sup>(49b) should not be confused with the acceptable (i)

(i) Ele...? ele partiu

where the first occurrence of *ele* is an 'afterthought' with an interrogative flavour and not a topic.

wh- questions – and relatives – IP is ‘dragged along as excess baggage’ – pied-piped – by the wh- phrase moving to the left periphery, as Ordoñez (1998) claims.

Naturally, the subject inversion structures studied here do not exhaust the cases of subject inversion in French and Portuguese, as (51) through (53) show:

- (51) a. Dans la forêt vivait un vieil ermite  
In the forest lived an old hermit  
b. À cette phrase correspond ce numéro  
To this sentence corresponds this number
- (52) a. Na floresta vivia um velho ermita  
(same as (51a))  
b. A esta frase corresponde este número  
(same as (51b))
- (53) a. Comeu o Pedro  
ate the Pedro  
‘Pedro has eaten’  
b. Telefonou o Pedro  
phoned the Pedro  
‘Pedro has phoned’  
c. Chegou o Pedro  
arrived the Pedro  
‘Pedro has arrived’  
d. Comeu a tarte o Pedro  
Ate the pie the Pedro  
‘Pedro ate the pie’  
e. Enviou o livro à Maria o Pedro  
Sent the book to Maria the Pedro  
‘Pedro sent Maria the book’

None of (53) has a well-formed counterpart in French, a well-known property generally seen as a consequence of the fact that French, unlike Portuguese, is not a pro drop language. As for (51), they are cases of locative inversion that both non pro drop languages and pro drop languages allow.

Whatever the comparative analysis one may adopt for such facts and differences one thing is clear: in each of these cases the postverbal subject is a focus, not a topic; the use we have made of the idea that SI structures involve Remnant IP movement as a side effect of Wh movement may ultimately give us a handle on this, but we shall have to leave the investigation of that possibility to future research.

## Acknowledgments

This article was written in French a few years ago, at a time when the final version of Kayne & Pollock (2001) hadn't come into existence and when much of our subsequent work on the left periphery of questions – cf. e.g. Ambar et al. (1998), Munaro, Poletto & Pollock (1999), Poletto & Pollock (2000), Pollock (2000) – hadn't even been started. In part because of that, the present work incorporates ideas and analyses that sometimes differ from what we have suggested in more recent work. However, the main ideas and concepts appealed to here we feel are basically right, which has prompted our decision to provide an English version with only minor modifications. Many thanks to João Peres and Rita Veloso for their help both as linguists and native speakers of Portuguese.

## References

- Ambar M. (1988), *Para uma sintaxe da inversão sujeito verbo em português*, doctoral dissertation, edição Estudos Linguísticos, Lisboa.
- Ambar M., H. Obenauer, I. Pereira & R. Veloso (1998), "The Internal Structure of wh-phrases in Portuguese and French", presentation at the "Structures formelles du langage" UPRESA 7023 round table, November 1998, Paris. Appears in Noam Chomsky's 70th birthday electronic Festschrift.
- Boivin M.-C. (2000), *Split Noun Phrases and the Theory of Case*, unpublished MIT dissertation, MIT.
- Cinque G. (1990), *Types of A-bar Dependencies*, MIT Press.
- Cornulier, B. de (1974), "Pourquoi et l'inversion du sujet non clitique", in C. Rohrer and N. Ruwet (eds.), *Actes du Colloque Franco-Allemand de Grammaire Transformationnelle*, vol. 1, Niemeyer, Tübingen, 139-163.
- Kampers-Mahne B. (1998), "'Je veux que parte Jean", a neglected construction", in Schegler A., B. Tranel & M. Uribe-Extebarria (eds.), *Romance Linguistics, Theoretical Perspectives*, John Benjamins, Amsterdam, 129-141.
- Kayne, R.S. (1984), *Connectedness and Binary Branching*, Foris, Dordrecht.
- Kayne, R.S. & J.-Y. Pollock (1978), "Stylistic Inversion, Successive Cyclicity and Move NP in French", *Linguistic Inquiry*, 9, 595-621.
- Kayne R.S & J.-Y. Pollock (1998), "New Thoughts on Stylistic Inversion", presentation at the Romance Inversion conference, Amsterdam, May 1998.
- Kayne R.S & J.-Y. Pollock (2001), "New Thoughts on Stylistic Inversion", in *Subject Inversion and the Theory of Universal Grammar*, A. Hulk & J-Y Pollock (eds.), Oxford University Press, New York.
- Munaro N., C. Poletto & J-Y Pollock (2001), "Eppur si muove! On Comparing French and Bellunese Wh-Movement", unpublished article, Padua & Amiens.
- Obenauer H. (1984), "On the identification of Empty Categories", *Linguistic Review*, 4.2.
- Ordoñez, F. (1998), "Post-Verbal Asymmetries in Spanish", *Natural Language and Linguistic Theory*, 16, 313-346.

- Poletto C. (2000), *The Higher Functional Field in the Northern Italian Dialects*, Oxford Studies in Comparative Syntax, Oxford University Press, New York.
- Poletto C. & J-Y Pollock (2000), "On the Left Periphery of some Romance Wh-questions", to appear in *The structure of CP and IP*, G. Cinque & L. Rizzi (eds.), Oxford University Press.
- Pollock J.-Y. (1986), "Sur la syntaxe de *en* et le paramètre du sujet nul", in M. Ronat & D. Couquaux (eds.), *La Grammaire Modulaire*, Editions de Minuit, Paris, 211-246.
- Pollock, J.-Y. (1998), "On the Syntax of Subnominal Clitics: Cliticization and Ellipsis," *Syntax*, 1, 300-330.
- Pollock J-Y (2000), "Subject clitics, Subject Clitic Inversion and Complex Inversion: Generalizing Remnant Movement to the Comp Area", unpublished article, Amiens, to appear in Syn.com.
- Rizzi L. (1982), *Issues in Italian Syntax*, Foris.
- Rizzi L. (1997), "The Fine Structure of the Left Periphery, in L. Haegeman (ed.), *Elements of Grammar*, Kluwer, 281-337.

Manuela Ambar  
*Faculdade de Letras da Universidade de Lisboa, Portugal*  
manuela.ambar@fc.ul.pt  
Jean-Yves Pollock  
*Université de Picardie à Amiens, France*  
jean-yves.pollock@wanadoo.fr