

The Structure of Split Interrogatives

Jorge López-Cortina
Georgetown University

1. Introduction

The purpose of this paper is to pursue a common analysis for a number of similar interrogative constructions that appear in several varieties of Spanish, as well as in other Romance languages spoken in the north of Spain. I will introduce the constructions, give a brief review of previous analyses (Lorenzo 1994, Camacho 2002), and present my own analysis, where the interpretation of the questions is explained by the position of their elements in the left periphery of the structure, in the line of work pursued by Rizzi (1997, 2001), Benincà (2001), Benincà and Poletto (2001), and others.

2. The construction

The general construction to be analyzed is exemplified by a particular type of wh-question, found in the Spanish spoken in Asturias, in which the wh-word *qué* can assume the value of any wh-word (*who, when, where, why, how*, etc.) when an appropriate answer is present in the sentence, on the right side and separated from the rest by comma intonation. I will refer to sentences that present this structure as split interrogatives. The complete sentence is, thus, a yes/no question in which a wh-word is present and co-refers with some other element. Approximate translations are given under the glosses¹.

- (1) a. Qué llegó , Juan?
what arrived-3SG, Juan?
'Who arrived? was it Juan?'
- b. Qué vienes , el sábado?
what come-2SG, the Saturday?
'When are you coming? on Saturday?'
- c. Qué vas , a Oviedo?
what go-2SG, to Oviedo?
'Where are you going? to Oviedo?'

The same questions are ungrammatical without the second part.

- (2) a. * Qué llegó?
- b. * Qué vienes?
- c. * Qué vas?

The second part, to the right of the comma intonation, can include multiple elements. In such cases the interrogative word refers to everything after the comma.

- (3) a. Qué llegó , Juan ayer?
 what arrived-3SG, Juan yesterday?
 b. Qué comió , pollo en casa?
 what ate-3SG, chicken at home?

It is more difficult to include elements between the verb and the intonation break marked by the comma, but it is easier to include a subject (4b) than a complement (4a).

- (4) a. * Qué llegó a Oviedo , Juan?
 what arrived-3SG to Oviedo , Juan?
 b. ? Qué llegó Juan , a Oviedo?
 what arrived-3SG Juan , to Oviedo?

These constructions allow for long extraction (5a), while they seem to be impossible in embedded clauses, as shown in (5b):

- (5) a. Qué dijiste que venía , Juan?
 what said-2SG that came-3SG, Juan?
 ‘Who did you say that was coming? was it John?’
 b. * Dijiste que qué venía , Juan?
 said-2SG that what came-3SG, Juan?

Similar constructions can be found among other non-Castilian northern Peninsular varieties of Spanish, as well as Chilean Spanish. They are also found in Catalan and in the Asturian language (one of the Leonese varieties, sometimes referred to as Bable). The examples in (4) are grammatical in many of these varieties. Chilean Spanish does not allow subjects to the right of the comma². Ecuadorian Spanish and some varieties of English present a construction where the wh-word remains *in situ*.

- | | | |
|--------|------------------------------------|----------------------------|
| (6) a. | Qué llegó Juan , a Oviedo? | (some Peninsular, Chilean) |
| | what arrived-3SG Juan , to Oviedo? | |
| b. | * Qué llegó , Juan? | (Chilean) |
| | what arrived-3SG Juan? | |
| c. | Vas qué, en tren? | (Ecuadorian) |
| | go-2SG what, in train? | |
| | ‘You go what, by train?’ | |
| d. | You are going what, by train? | (English) |

Finally, Catalan also presents a different construction with a *que* that could be an overt complementizer or a question marker similar to Japanese *ka* (7a). This is particularly interesting because it shows very clearly the difference in intonation and meaning between the split interrogatives and other yes/no questions.

- (7) a. Que vas en tren? (Catalan)
 that go-2SG in train
 ‘Are you going by train?’
 b. Què vas , en tren? (Catalan)
 what go-2SG, in train
 ‘How are you going? by train?’

When the interrogative word is more specific than the neutral *qué*, or *qué* is the interrogative word to be expected, the construction has frequently been interpreted as the compound or juxtaposition of two interrogative sentences (Py 1971, Lorenzo 1994). Vigara Tauste (1992) calls these constructions expressive anticipations (*anticipación expresiva*) and assumes that the first part is not an independent interrogative; she refers to that first part as a doubled element (*elemento geminado*), but presents no further analysis; Camacho (2002) calls the constructions wh-doubling and notes differences between them and other interrogatives. His analysis is presented in the following section.

3. Previous analyses

3.1. Camacho's (2002) Wh-Doubling

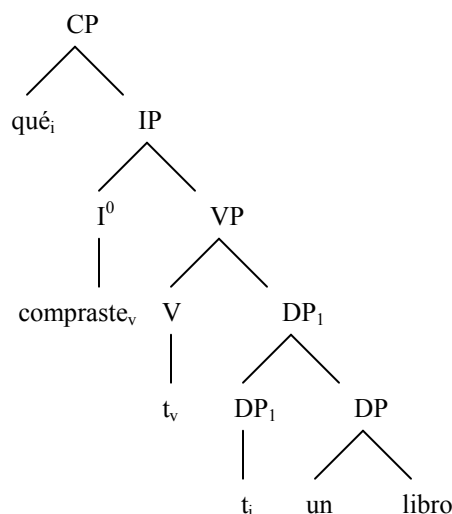
Camacho (2002) analyzes the cases where the wh-word in the first part of the question is *qué* and the second part is a DP (as *un libro* ‘a book’ in 8a) or a CP. In such constructions no interrogative other than *qué* is expected, unlike in (1), and therefore the first part of the question would be grammatical by itself. The wh-word and the DP or CP that constitute the second part of the question are related, and they are both interpreted as a complement of the verb *comprar* ‘to buy’ in (8a). Locality effects are also observed, as shown in (8b):

- (8) a. Qué compraste, un libro?
 what buy-PAST-2SG, a book?
 ‘What did you buy? a book?’
 b. * Qué te molesta el hecho de que comprara , un libro?
 what CL-2SG bothers the fact of that buy-PAST-2SG, a book?
 ‘What does it bother you the fact that he bought, a book?’

Camacho proposes that the DP or CP is adjoined to the trace of the wh-word (9)³. This would explain the common argument position and the particular intonation of the construction.

- (9) a. Qué compraste, un libro?
 what buy-PAST-2SG, a book?
 ‘what did you buy? a book?’

b.



Camacho takes locality effects, as in (8b), as an indication of a relationship between the *wh*-word and a position inside the second part of the question. The ungrammaticality of (8b), however, follows from the extraction of the *wh*-word and does not necessarily prove the connection between the *wh*-word and the DP in sentence-final position. Notice that without the DP the sentence is still ungrammatical, as in (10):

- (10) * Qué te molesta el hecho de que comprara?
 what CL-2SG bothers the fact of that buy-past-2SG?
 ‘What does it bother you the fact that he bought?’

Locality effects still have to be explained if we assume that in these cases there is a single construction with two interdependent parts, rather than two juxtaposed questions. A strong argument in favor of the single construction approach would be the existence of cases where the parts are not grammatical if taken separately, as in (2).

Camacho’s (2002) proposal captures the intuition both that both parts of the question conform a single structure, and that the *wh*-word and the second part are somehow related. It does not explain the cases where intervening elements render the construction ungrammatical, like those in (4a), and it leaves open the issue of the nature of the *wh*-word *qué*, which would undergo *wh*-movement but would not be a [+*wh*] operator. More importantly, it does not provide a

motivation for the doubling. I will address these issues in section 5. In the following section I present Lorenzo's (1994) alternative analysis.

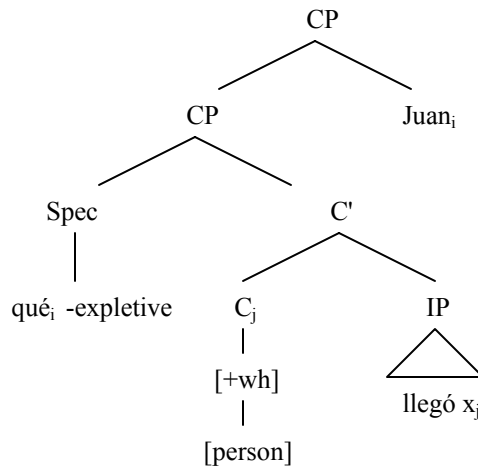
3.2. Lorenzo's (1994) analysis

Lorenzo (1994) studies the cases described in (1), repeated here as (12), where the first part of the question is not viable as an independent question. He proposes the analysis illustrated in (12b), where the focus on the right is dislocated (as opposed to topicalized; the terminology corresponds to that in Rivero 1980), and therefore base-generated. Lorenzo considers the interrogative in Spec CP to be an expletive, in spite of apparent evidence for movement, like the presence of Subjacency effects (11), which he attributes to the need to establish a chain between an empty operator in C^0 and an empty category in VP.

- (11) a. Qué debíamos alegrar, a Juan?
 what should-1PL cheer-up , to Juan?
 'Whom were we supposed to cheer up? Juan?'
 b. * Qué te empeñaste cuándo debíamos alegrar, a Juan?
 what CL-2SG insist-2SG when should-1PL cheer-up , to Juan?
 'Whom did you insist when we were supposed to cheer up? Juan?'

- (12) a. Qué llegó, Juan?

b.



Lorenzo's argument in favor of dislocation is based on the contrast in grammaticality between (1a) and (4a), repeated here as (13).

- (13) a. Qué llegó e_i , Juan_i?
 what arrived-3SG, Juan?
 b. * Qué llegó e_i a Oviedo , Juan_i?
 what arrived-3SG to Oviedo , Juan?

According to Lorenzo, (13b) is ungrammatical because the empty category is not properly governed (Lasnik and Saito 1992). The verb *llegó* 'arrived' from (13b) can not properly govern the empty category because it is already involved in a thematic coindexation with a complement, in contrast with the grammatical (13a), in which there is no complement. An analysis of (13b) where the element to the right had undergone movement would imply the existence of a trace of that movement within the IP. This trace would be properly governed, thus licensing the chain and making (13b) grammatical. Therefore, Lorenzo rejects the movement analysis in favor of the dislocation analysis.

The most apparent problem with Lorenzo's structure is that the expletive and its associate c-command each other. Lorenzo notes this and sketches an alternative structure within Kayne's (1994) framework, where the construction would be derived by adjoining an interrogative sentence to a "reduced" sentence structure through an empty functional head, as in (14).

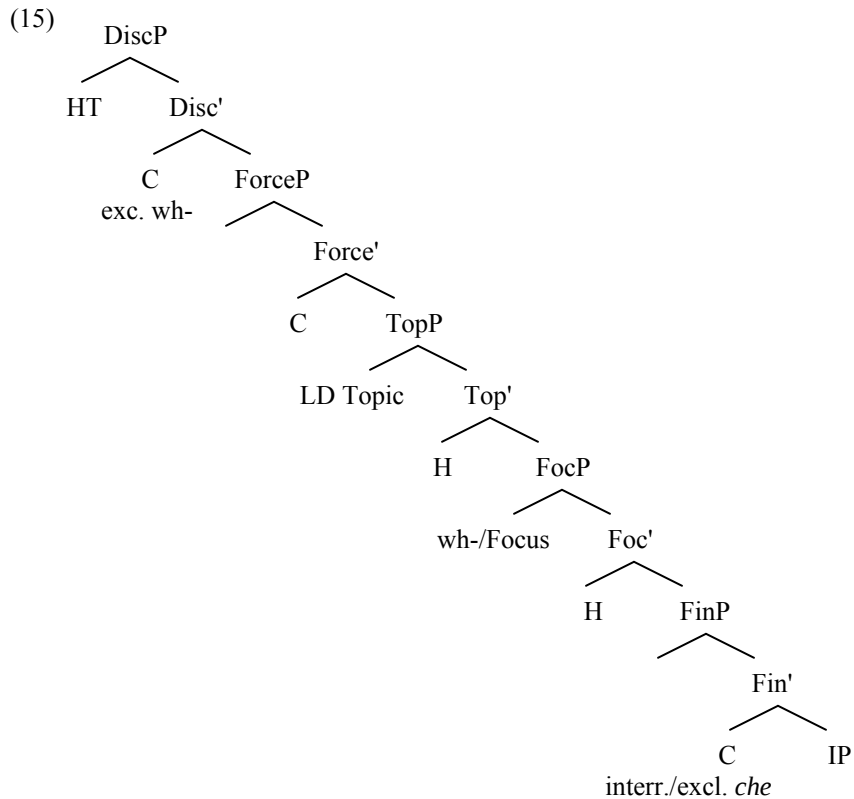
(14) [Qué llegó? [X⁰ [Juan (llegó)?]]]

There are a number of issues that, in my view, require further explanation. For instance, (i) How exactly would the limitations of the construction be accounted for in a system that excludes right-adjunction? (ii) What exactly is that *qué* expletive? Where does it come from? (iii) How can dialectal variation be accounted for? The examples that Lorenzo (1994) gives as key for his structure are precisely the ones that vary in acceptability from one dialect to another. (iv) Why does this construction exist in the first place? i.e., what can be expressed with this construction that can not be expressed with the other, more frequently analyzed, interrogative constructions?

This last issue is perhaps best approached within the current study of the left periphery, where information structure issues are articulated in syntactic terms. I will now introduce Benincà's (2001) proposal of structure of the left periphery, which will serve as framework for my own analysis of split interrogatives.

4. Benincà's left periphery

The structure of the CP layer (the so-called left periphery) as presented by Benincà (2001) is given in (15):



This structure has a number of advantages over previous proposals. In particular, the new Disc(ourse)P(hrase) provides the space for a base-generated thematic element, realized in Italian as a Hanging Topic (HT).

Following Rizzi's (1997) formulation, the two informational articulations of the clause (topic-comment, focus-presupposition) are represented in the syntactic structure by functional projections, but Benincà's structure eliminates both the recurrence of the Top(ic)P(hrase) and the lack of a definite position for TopP with respect to FocP that characterized it in Rizzi's proposed structure. The problem of multiplicity of phrases in the left periphery is transferred to Spec, FocP, which Benincà considers able to host multiple elements, of which only the first one would receive focalized intonation. Benincà identifies Spec, FocP with the wh- landing site⁴.

In what follows, I will accept Benincà's structure as a tested framework for several Romance languages (in addition to Italian, Benincà draws evidence from French and Paduan) and will not discuss here the peculiarities of Spanish constructions other than the one that is the subject of this paper.

5. Proposed analysis

Let us follow Benincà and Poletto's (2001) suggestion that Focus is not a single projection but itself a sub-field, and separate a projection containing Focus' most distinct element, the interrogative *wh*-landing site. For the purposes of this paper, let us call that projection WhP(hrase). This phrase would go immediately below FocP (see (16): while examples of a focalized element to the left of an interrogative might not be obvious (16a), finding it to the right is simply impossible (16b). Such examples can only be interpreted as rhetorical interrogatives (16c), where the *wh*-word would be up in the exclamative *wh*-position⁵).

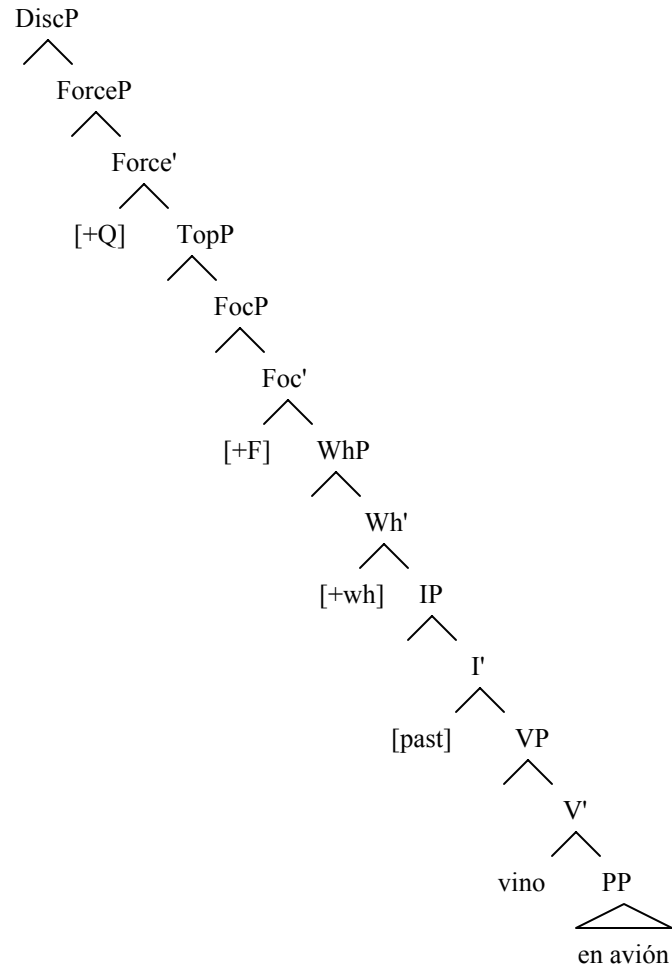
- (16) a. – Voy a estudiar mañana.
 go-1SG to study tomorrow
 'I am going to study tomorrow'
 – HOY, qué vas a hacer?
 today, what go-2SG to do
 'TODAY, what are you going to do?'
 b. – Voy a estudiar mañana.
 go-1SG to study tomorrow
 'I am going to study tomorrow'
 – * Qué, HOY, vas a hacer?
 what, today, go-2SG to do
 'What, TODAY, are you going to do?'
 c. What, in these modern times, can we expect of the youth?

Consider now a simple sentence like (17). I will not consider any of the elements to be generated in their final positions outside of the VP⁶. The sentence will be derived on the basis of the features present in the CP layer.

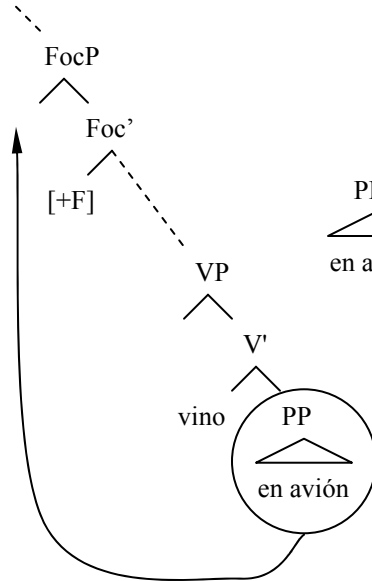
I propose that the final output is the result of three movement operations, detailed in (18a) through (18d).

- (17) a. Qué vino , en avión?
 what came-3SG, in plane
 'How did he come? By plane?'

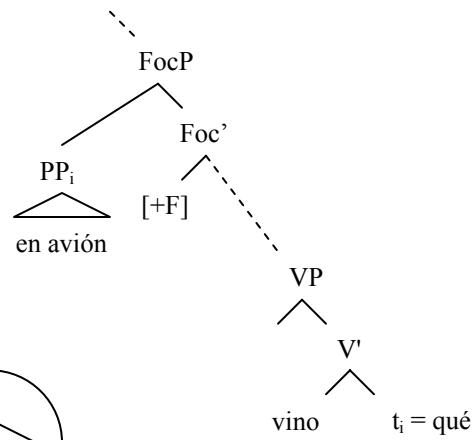
(17) b.



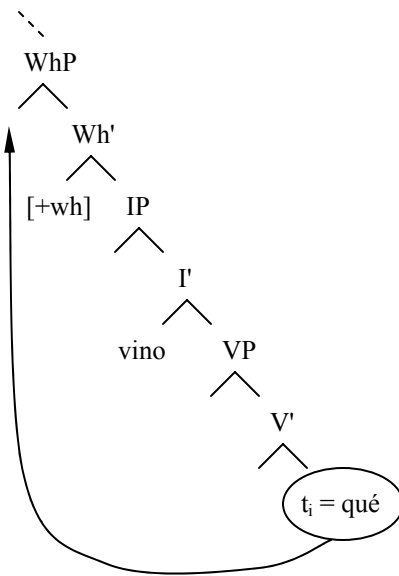
(18) a.



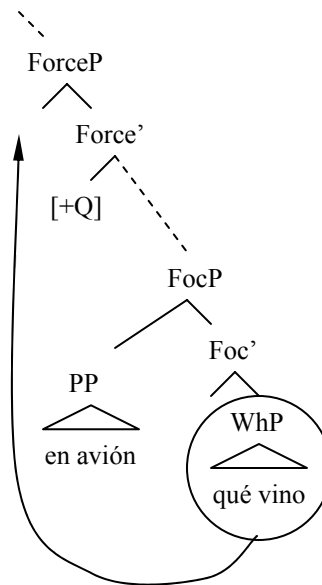
b.



c.



d.



(18a, b) illustrate the focalization of the element that will appear to the right of the comma intonation. This movement leaves a trace, which will be spelled out as *qué*. (18c) shows the wh-movement of the trace left by the focalization. This movement does not take place in Ecuadorian Spanish, where the *qué* is left *in situ*. In Chilean Spanish it is optional. Finally, (18d) shows the movement to Spec ForceP of the entire WhP. This is what makes the sentence a question, as opposed to a sentence where an element has simply been focalized. The idea of a [+Q] feature has been around at least since Baker (1970), and it would be independent from the [+wh] feature. Both features are clearly separated in the case of wh-relatives, which are [+wh] but not [+Q]⁷.

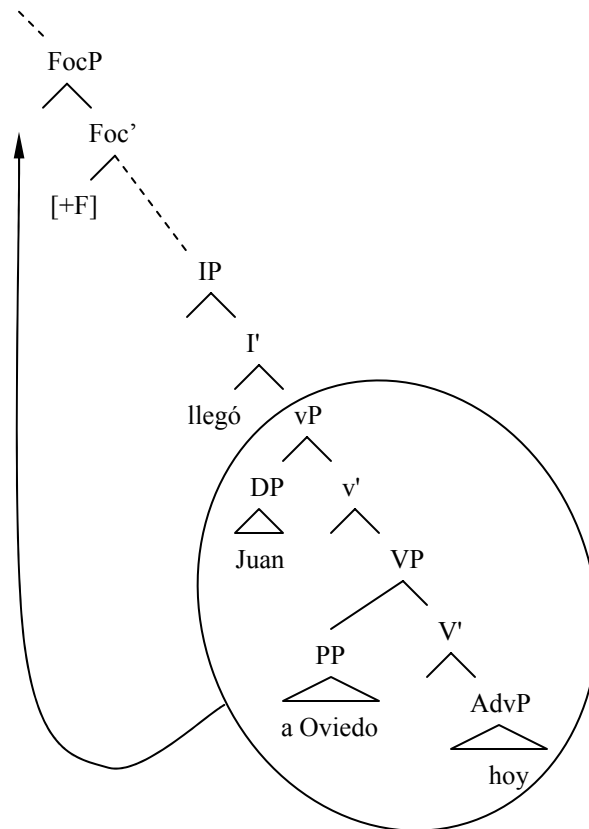
I propose that the *qué* that appears in the structure is an overt trace left by the focalization. The reason the trace is realized overtly is because after the movement shown in (18d) it is left unbound, and therefore it would violate ECP. Phonetically realized traces have been proposed for Swedish (Engdahl 1985), a language where resumptive pronouns appear in wh-questions and can license parasitic gaps. An anonymous reviewer points out that resumptives spelled out as wh- are unheard of. In fact, Swedish resumptive pronouns that are variables in wh-questions are not spelled out as wh-, in spite of being traces of wh-phrases (in Engdahl's analysis). Here I present the inverse situation, where the trace retains the [+wh] feature. It might be plausible that a [+wh] resumptive is spelled out as wh- in such cases. However, this still leaves some questions open, particularly about the compatibility of the analysis with the copy theory of movement.

As for the wh-movement (18c), there are dialects both in which *qué* moves and in which *qué* stays *in situ*, so I will not speculate here about the exact characteristics of its movement.

Let us go back now to the question about the reason for the existence of this construction, i.e., what this construction expresses that other interrogative constructions do not. This kind of question is used when the speaker wishes to obtain confirmation for some information she has. This construction allows for the unambiguous expression of the fact that, in spite of the existence of a phrase that completes the information, the question posed is not a yes/no question, but rather one about a particular element in the sentence. In this analysis, this is reflected as the confluence of focus and interrogative features in the same phrase.

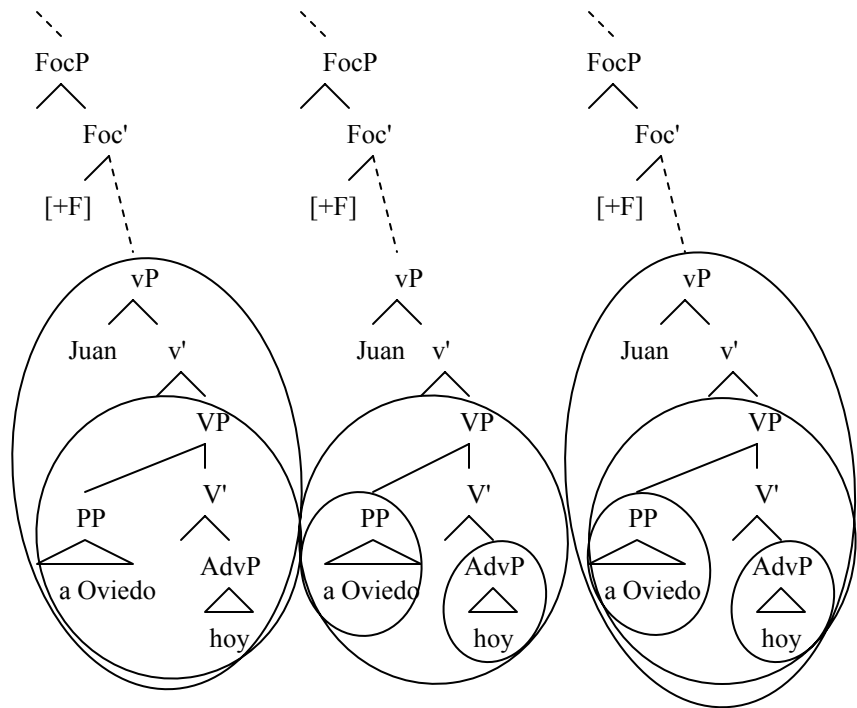
The analysis of multiple elements to the right of the comma intonation is straightforward once Larsonian VP shells (Larson 1988) are assumed: the entire vP is focalized once the verb has raised to I.

- (19) Qué llegó , Juan a Oviedo hoy?
 what arrived-3SG , Juan to Oviedo today?



Dialectal variation would follow from the different possibilities for the focalization. Asturian Spanish allows focalization only of verbal projections, either including or excluding the subject. Other northern Peninsular Spanish varieties allow focalization of any given projection below IP, and Chilean Spanish allows focalization of any projection below the one that contains the subject⁸.

(20)



Asturian

Chilean

Other Peninsular

Qué llegó, Juan a Oviedo hoy?	* Qué llegó, Juan a Oviedo hoy?	Qué llegó, Juan a Oviedo hoy?
Qué llegó Juan, a Oviedo hoy?	Qué llegó Juan, a Oviedo hoy?	Qué llegó Juan, a Oviedo hoy?
* Qué llegó Juan hoy, a Oviedo?	Qué llegó Juan hoy, a Oviedo?	Qué llegó Juan hoy, a Oviedo?
* Qué llegó Juan a Oviedo, hoy?	Qué llegó Juan a Oviedo, hoy?	Qué llegó Juan a Oviedo, hoy?
* Qué llegó a Oviedo hoy, Juan?	* Qué llegó a Oviedo hoy, Juan?	Qué llegó a Oviedo hoy, Juan?

Evidence for this analysis can be found in sentences like (21a) (from López-Cortina 1997) where the focalized element includes everything below IP (but not IP).

- (21) a. Qué está llegando el tren?
 what is, arriving the train?
 ‘What is going on? Is the train arriving?’
 b. Qué está llegando, el tren?
 what is arriving, the train?
 ‘What is arriving? Is it the train?’

It could be hypothesized that the IP should also be susceptible to focalization. And in fact, examples like standard Spanish in (22) seem to support that hypothesis:

- (22) Qué, llegó Juan?
 what, arrived-3SG Juan?
 ‘Well, did Juan arrive?’

6. Conclusions and extensions

The constructions analyzed here offer an interesting testing ground to study the relationship between syntax and the articulation of Topic, Focus and Interrogation. I have given an analysis that interprets the default *wh*-word in these constructions as a trace left by focalization, and the final word order as a result of the coexistence of focalization and *wh*- features. This analysis has some advantages, among them a) giving a uniform account for a variety of dialects, b) being compatible with independently motivated analyses of the left periphery, and c) motivating movement with features that match the actual meaning of the sentence.

Further research should address the status of the diverse topicalization/dislocation constructions that are particular to Spanish within the proposed hypotheses for the CP layer. Another area open to research is the existence of similar restrictions on focalization in non-interrogative sentences.

It would be interesting to test whether a similar analysis would apply to cases like (23), even though they do not present the same kind of restrictions.

- (23) a. Quién llegó a Oviedo, Juan?
 who arrived-3SG to Oviedo, Juan?
 ‘Who arrived to Oviedo? Was it Juan?’
 b. Dónde llegó Juan, a Oviedo?
 where arrived-3SG Juan, to Oviedo?
 ‘Where did Juan arrive? Was it Oviedo?’

Notes

- * I would like to thank Héctor Campos, Raffaella Zanuttini, and about twenty native speakers of different dialects for their support, insights, and patience, as well as two anonymous reviewers for their useful comments. All mistakes remain my own.
1. Throughout the paper I will omit the reverse question mark of standard Spanish orthography, in order to avoid confusion with the ? used to indicate marginality.
 2. Chilean Spanish presents subjects to the right of the comma if the wh-word is left in situ, as in *Llegó qué, Juan?* (Héctor Campos, p.c.)
 3. Alternatively, Camacho presents another analysis within Kayne's (1994) framework (where there is no adjunction). In that analysis, the trace and the DP/CP would be part of a small clause complement, but he notes that there are several problems with it, for instance the lack of other cases of small clauses inside DP in Spanish.
 4. This could go against the analysis of Rivero (1980), who showed how the movement that she called Topicalization (and which is indistinguishable from Rizzi-style Focalization except for the lack of contrastive intonation) is different from wh-movement in that it is not sensitive to wh-islands.
 5. An anonymous reviewer notes that examples like (16c) may appear embedded, and this would suggest that the wh- element could not be in Spec, DiscP after all. Independently of what the best analysis for such examples might be, the relevant fact for our purposes is that the reply in (16b) is not acceptable if *hoy* 'today' is focalized.
 6. For the sake of clarity, I omit the subject from the trees in (17) and (18).
 7. I would like to thank Héctor Campos and an anonymous reviewer for this observation.
 8. Paula Kempchinsky (p.c.) suggested, as an alternative explanation for the Chilean data, that the subject could be occupying a different position in Chilean than it does in the other dialects. This would reduce the dialectal variation of these structures. Independent evidence of a different position for Chilean subjects should be found before deciding in favor of either approach.

References

- Baker, Carl L. 1970. Notes on the Description of English Questions: The Role of an Abstract Question Morpheme. *Foundations of Language* 6. 197-219.
- Benincà, Paola. 2001. The position of Topic and Focus in the left periphery. *Current Studies in Italian Syntax offered to Lorenzo Renzi*, ed. by Guglielmo Cinque and Giampaolo Salvi. 39-64. Amsterdam; New York : Elsevier.
- Benincà, Paola, and Cecilia Poletto. 2001. Topic, Focus and V2: defining the CP sublayers. Talk at Georgetown University.
- Camacho, José. 2002. Wh-Doubling: Implications for the Syntax of Wh-Movement, *Linguistic Inquiry* 33, 1. 157-164.
- Engdahl, Elisabet. 1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23. 3-44.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.
- Larson, Richard. 1988. On the Double Object Construction, *Linguistic Inquiry* 19, 3. 335-391.
- Lasnik, Howard, and M. Saito. 1992. *Move α : Conditions on Its Application and Output*, Cambridge: MIT Press.

- López-Cortina, Jorge. 1997. Acerca de algunas construcciones interrogativas del español. Paper presented at the 47th MIFLC, East Carolina University.
- Lorenzo, Guillermo. 1994. Qué expletivo en preguntas dislocadas. *Archivum Ovetensis* XLIV-XLV, 1. 423-446.
- Py, Bernard. 1971. La interrogación en el español hablado de Madrid. Bruxelles: Aimav.
- Rivero, Maria Luisa. 1980. On Left-Dislocation and Topicalization in Spanish, *Linguistic Inquiry* 11, 2. 363-393.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of Grammar*, ed. by Liliane Haegeman. 281-337. Dordrecht: Kluwer.
- Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. *Current Studies in Italian Syntax offered to Lorenzo Renzi*, ed. by Guglielmo Cinque and Giampaolo Salvi. 287-296. Amsterdam; New York : Elsevier.
- Vigara Tauste, Ana María. 1992. *Morfosintaxis del español coloquial*. Madrid: Gredos