

The Cl(itic) Projection in Questions^{*}

Maria Luisa Zubizarreta

University of Southern California. Department of Linguistics
Los Angeles, California 90089-1693. USA
zubizarr@rcf.usc.edu

Received: December 13th 1998

Accepted: March 17th 1999

Abstract

Metadata, citation and similar papers at core.ac.uk

provided

wh-phrase and the verb. It is suggested that the difference stems from a fundamental structural difference between the two types of languages. Romance, unlike English, projects an abstract *Cl(itic)* category above TP. *Cl* is an abstract operator whose function is to «externalize» an argument of a verb *v* with respect to the tense associated with *v*. *Cl* is identified by rich agreement or a clitic morpheme, which accounts for the presence of such a category in Romance and its absence in English. It is then argued that the projection of a *Cl*-operator between CP and IP in wh-questions gives rise to a minimality effect. The appendix deals with some cases of clitic-doubling of non-specific wh-phrases in modal contexts and suggests that the clitic (associated to *Cl*) functions as an E-type pronoun in such cases.

Key words: wh-questions, clitics, syntax.

Resum. *La projecció Cl(ític) a les interrogatives*

En aquest article s'intenta explicar una restricció força coneguda, molt freqüent a les llengües romàniques, que fa que el subjecte a les interrogatives-qu no pugui intervenir entre un mot-qu i el verb. Es suggereix que la diferència prové d'una diferència estructural fonamental entre dos tipus de llengües. A les llengües romàniques, a diferència de l'anglès, es projecta una categoria abstracta *Cl(ític)* per sobre del ST. El *Cl* és un operador abstracte la funció del qual és la d'«externalitzar» un argument d'un verb *v* respecte al temps associat amb *v*. El *Cl* s'identifica per la concordança rica o un morfema clíctic, que dona compte de la presència d'aquesta categoria a les llengües romàniques i la seva absència en l'anglès. Es defensa que la projecció d'un operador-*Cl* entre el SC i el SI a les interrogatives-qu dona lloc a un efecte de minimitat. A l'apèndix es tracten alguns casos de doblament de clíctic de mots-qu no específics en contextos modals i s'apunta que el clíctic (associat amb *Cl*) funciona com un pronom de tipus-E en aquests casos.

Paraules clau: interrogatives-qu, clíctics, sintaxi.

^{*} This article (with the exception of the appendix) will also be published in the Proceedings of the Conference «On Romance Inversion» (1998), edited by Hulk, A.; Pollock, J.-Y., Oxford University Press. I would like to thank Aafke Hulk for very useful comments and discussion.

Table of Contents

1. Introduction

There is a wellknown constraint, pervasive in the Romance languages, which disallows the subject to intervene between a fronted *wh*-phrase and the verb. We illustrate this constraint with examples from Standard Spanish, French, European Portuguese, and Italian:

- (1) a. *Qué Pedro compró? (Spanish)
 what Pedro bought
 b. *Que Pierre a acheté? (French)
 c. *Que o Pedro comprou? (EPortuguese)
 d. *Che Piero ha comprato? (Italian)

The existence of such a constraint can also be appreciated in constructions like French complex inversion (where V-to-C has applied); it minimally contrasts with English, where the constraint does not apply:¹

- (2) a. A qui a-t-il (*Pierre) parlé?
 to whom has-nom.cl. (Pierre) spoken
 b. To whom did Peter speak?

In Zubizarreta (1998), I proposed that the preverbal subject in Spanish is always associated with a discourse-related feature («topic» or «focus» or «emphasis»). Two further assumptions will then conspire to account for the constraint on preverbal subjects in interrogatives: 1) the «topic» feature is always projected above the «focus» feature at the left-edge of the clause and 2) the *wh*-phrase is associated with the «focus» feature. It therefore follows that a «topic» subject must precede the *wh*-phrase when they cooccur at the left-edge.

1. As Poletto (in press) argues, the restriction on the position of the preverbal subject in interrogatives is independent from the V-to-C parameter. This is shown by examples in Northern Italian dialects such as the following, where no V-to-C has occurred but the restriction on the preverbal subject is still present.

- (i) i ho domanda kol ke (??Mario) l'ha fat
 to-him have asked what that (Mario) he has done

Suñer (1994) makes the same point for Spanish.

There are various problems with such analysis. One of them is lack of generality. The constraint is also active in French, a language which lacks a «focus» projection at the left-edge of the clause. It is furthermore empirically insufficient. While the constraint on preverbal subjects holds systematically in root clauses in most Romance languages, there is variability in embedded clauses. While it applies in embedded interrogatives in Spanish and Italian, this is not the case in French and European Portuguese; e.g., cf. the Spanish/French contrast below:

- (3) a. *Me pregunto a quién Juan busca.
 (I) wonder who Juan is looking for
 b. Je me demande qui Jean cherche.

Another fact that the above analysis does not readily account for is the contrast between bare and nonbare *wh*-phrases. While bare *wh*-phrases activate the constraint on preverbal subjects, non-bare *wh*-phrases like *which N* and *how many N* do not (cf. Ambar (1988), Pollock et al. (1998)):

- (4) a. (?) Cuál libro Pedro compró?
 Which book Pedro bought
 b. (?) Quel livre Pierre a acheté
 c. Que livro o Pedro comprou?
 (5) a. A cuánta gente María invitó a la fiesta?
 How many people María invited to the party
 b. Combien de gens Marie a invité à la fête?

In this article, I will articulate a different proposal. I will suggest that there are two basic ingredients which interact to account for the unexceptional application of the constraint in most Romance languages and its variability in embedded interrogatives. I) The existence of the so-called clitic-left dislocation construction pervasive in the Romance languages but inexistent in English. We assume (in the spirit of Iatridou (1990)) that the existence of the clitic-left dislocation (CLLD) construction in the Romance languages indicates that the predicate-argument structure may be syntactically articulated in these languages in a different way than in English. While the verbal arguments in English are systematically projected within the VP, in the Romance languages under discussion they may be merged with an abstract (λ) operator generated above TP. Such operator in turn binds an argument-variable within the VP. We will refer to it as the CI operator because the argument-variable it binds is typically identified by a clitic or «strong» inflectional element.² See section 2. II) The Q feature in the left edge of

2. Note that within this view the topic nature of the «clitic-left dislocated» argument is entirely tangential to its grammatical properties. The discourse property of «topichood» is extrinsic to the grammatical properties of the construction.

the CP field (which expresses the interrogative force of the sentence) must be phonologically lexicalized, but there are more options available for lexicalizing a Q in embedded than in matrix interrogatives (see section 3). When the Q-feature binds a variable, it acquires the status of a syntactic operator and interacts with the CI operator giving rise to a minimality effect from which the preverbal subject constraint follows immediately (see section 4).

2. The CI Projection

What is this CI-projection? It is neither a scope-position associated with a +specificity feature (as suggested by Sportiche (1992, 1998)) nor an Agr position (as suggested by Franco (1993)). We suggest instead that:

- (6) CI is an abstract operator whose function is to «externalize» an argument of a verb v with respect to the tense associated with v .

Externalization of an argument via a CI operator thus gives rise to the following type of structure:

- (7) [DP_i [CI_i [T [_{VP} ... [V e_i ...]]]]]

The proposal is then that there is more than one way in which the predicate-argument relation may be syntactically instantiated:

- (8) a. In the familiar way, via projection within VP, in which case the DP argument is merged with V or with a projection of V: [V DP] or [DP [V ...]]
- b. The DP argument is merged with the projection of a CI operator that binds an argument variable within the VP. The DP is coindexed with CI via Spec/head agreement and therefore with the argument variable that CI binds: [DP_i [CI_i ... [_{VP} ... e_i ...]]]

We may refer to the structure obtained via projection of DP arguments within VP as the lexical or small proposition and to the structure obtained via merging of DP with the CI operator as the syntactic or large proposition. (8b) has as its predecessor the analysis of the CLLD construction put forth by Iatridou (1990). Although the details of the proposal put forth here differs in important respects from that of Iatridou's, they share a basic insight, namely that the predicate-argument relations in the CLLD construction is articulated in a fundamentally different way from other constructions.

The most convincing argument in favor of (8b) are the binding facts first reported in Zubizarreta (1993) for Spanish and reproduced in other languages with a CLLD construction such as Arabic (Aoun and Benmamoun (1998)) and Greek (P. Schneider-Zioga, p.c.). See also Cecchetto (1995) and Guasti (1996) on Italian.

This data shows that there is a pre-verbal/post-verbal subject asymmetry in cases where the fronted accusative topic (which contains the intended bindee) and the subject (which contains the intended binder) originate in the same clause.³

- (9) a. A su_i hijo, cada madre_i deberá acompañarlo_i el primer
acc his child each mother must accompany-acc.cl the first
día de escuela.
day of school
- b. ?*El primer día de escuela, a su_i hijo deberá acompañarlo_i
The first day of school acc his son will accompany-acc.cl
cada madre_i
each mother
- (10) a. A su_i propio hijo, ningún padre lo_i quiere castigar.
acc his own child no father acc.cl wants to punish
'No father wants to punish his own child.'
- b. ?*A su_i propio hijo no lo_i quiere castigar ningún padre_i
acc his own child not acc.cl wants to punish no father
'No father wants to punish his own child.'

No asymmetry is found when the binder is the matrix subject and the bindee is contained within a fronted accusative object that originates in the embedded clause. See (11).

- (11) a. A su_i hijo, ninguna madre_i debió aceptar que
acc his son, no mother should have accepted that
se lo regañe.
indef. subj.cl acc.cl scold
- b. A su_i hijo, no debió aceptar ninguna madre_i que
acc his son, not should have accepted any mother that
se lo regañe.
indef. subj.cl acc.cl scold
'No mother should have accepted that someone scold his son.'

The above data suggests that there is reconstruction of the fronted object to a position within the clause that it originates, but that this position is not within the

3. It has been brought to my attention by Xavier Villalba that for some speakers the binding relation in (9a) and (10a) is facilitated by the presence of a modal. For such speakers, the availability of a bound reading decreases in the absence of a modal, as in the following example:

(i) A su_i hijo, cada madre_i lo acompañó el primer día de escuela
acc his child, each mother acc.cl accompanied the first day of school

This observation, if correct, deserves further investigation.

VP (cf. the illformedness of (9b) and (10b)). The fronted object may be reconstructed to a position above the VP and below the preverbal subject of the clause where it originates (cf. the wellformedness of (9a) and (10a)). The fact that there is no reconstruction of the fronted object within the VP of the clause it originates suggests that the fronted object is inserted higher up in the structure, i.e. in the Spec of Cl (cf. Zubizarreta (1993, 1998)). This means that there is no copy of the fronted object in the canonical object position within the VP; the binding facts then follow immediately.

The proposal we are putting forth here may be summarized as follows:

- (12) In languages with a nominal «clitic» or «strong agr» system, a direct argument of the verb (nominative or accusative) is «externalized» by merging the argument with Cl (located above T) rather than by merging the argument with V (or a projection of V) and then moving it outside the scope of T.

It follows from (12) that the preverbal subject in languages with a CLLD construction of the sort under discussion is also generated in Spec of Cl.⁴ This means that in sentences like (9a) and (10a), there are two Cls: one externalizes the nominative argument and the other one externalizes the accusative argument. As the binding facts indicate, in these examples the nom-related Cl is generated higher than the acc-related Cl. The dislocated object may therefore «reconstruct» to a position within the scope of the nominative subject. But it is also possible for the acc-related Cl to be generated higher than the nom-related Cl. As predicted, in such cases the fronted object may bind into the fronted subject:

- (13) a. A cada niño, su madre lo acompañará el
 acc each child (obj), his mother (subj) acc. cl accompany-fut the
 primer día de escuela.
 first day of school
- b. A ninguno de estos niños, su padre lo
 acc none of these children (obj), his father (subj) acc. cl
 quiso acompañar.
 wanted to accompany

Hence, the sentences in (9a) and (10a) have the structure in (14a) and the sentences in (13) have the structure in (14b).⁵

4. The hypothesis that preverbal subjects in some of the Romance languages does not originate within the VP is not a novel one (e.g. Barbosa (1995)). Here we are generalizing the assumption to all languages with a CLLD construction.
5. A question that arises at this point is why the structures in (14) do not violate minimality. This question is particularly relevant in view of the fact that, as we shall see in section 3, when Cl interacts with Q, it does give rise to minimality effects. It would seem that two identical operators (i.e. two operators with exactly the same function) do not interfere with one another, but two distinct operators may do so. See section 4.3 for further discussion.

- (14) a. $[DP_j [Cl_i] [DP_i [Cl_i] [T [_{VP} [e_j] [V [e_i] \dots$
 b. $[DP_i [Cl_i] [DP_j [Cl_j] [T [_{VP} e_j] [V [e_i] \dots$

Let us turn back to the generalization in (12). Why is there a connection between the existence of Cl and of morphological clitics or strong agreement affixes? The answer is that the abstract category Cl must be morphologically identified in the following manner:

- (15) Morphological identification of Cl is achieved by some nominal morpheme with phi-feature contained within the local T. This may be a clitic or a «strong» agreement affix.

The clitic or verbal agreement affix is of course related to the DP in Spec of Cl. In the case of the clitic, this relation is captured by the fact that a nominal clitic is nothing else than a copy of the D and phi-features of the argument-variable bound to the DP. In the case of the agreement affix, this relation is captured by the Spec/head relation that the argument-variable bound to the DP bears with the verb at some point in the derivation. See below for further details.

We can now raise the question of what is the relation between the type of morphological identification required by Cl (as summarized in (15)) and the one imposed by the *Functional Projection Activation Principle* (FPAP) put forth by Speas (1994), Koopman (1997), Poletto (1998), and others:

- (16) *Functional Projection Activation Principle*: A feature in the head of a functional projection is activated iff the appropriate morphological material occupies either the head or the Spec of that projection at some point in the derivation.

Note that the requirements on the morphological identification of Cl is stricter than the one imposed by the FPAP. In effect, a DP argument in the Spec of Cl is insufficient to identify Cl. The reason is that Cl does not encode a (semantic) feature like other functional categories. Its role is purely formal, namely that of establishing a connection between a DP in its Spec and an argument-variable of the verb. More precisely, Cl serves the role of externalizing a DP argument with respect to the T associated with V. Some morphological nominal element on such a T is required to identify Cl. Another way of putting it is that the position of syntactic activity of Cl and the position of its feature-content are dissociated; its feature-content is not under Cl as one might expect. Rather, it is contained within the T node immediately below it.

The acc-related Cl is typically identified by an accusative clitic. There is no other option in the Romance languages because these languages lack morphological object agreement on the verb. The accusative clitic is a copy of the D and phi-features of the accusative argument-variable. It left-adjoins to V prior to the application of V-to-T (I assume with Kayne (1994) that adjunction is always leftward):

- (17) $[DP_i [Cl_i [[_{V_j} \text{acc cl}_i [v+agr]] T] [_{VP} \dots [V_j e_i] \dots$

The identification of the nom-related CI is not uniform across the Romance languages. Some languages (like many of the Northern Italian dialects) have a nominative clitic which is systematically used to identify the nom-associated CI. The nominative clitic is a copy of the D and phi-features of the nominative argument-variable. Since the nom argument-variable is located above the verb in the verbal projection, its copy (i.e. the nom clitic) cannot left-adjoin to the verb. Instead it left-adjoints to the T to which the verb has been adjoined. (As we shall see in the next section, this difference between nominative and accusative clitics is important because it accounts for why accusative clitics are always proclitic, but nominative clitics are enclitics when V-to-C applies).

(18) $[DP_i [CI_i [nom\ cl_i [[v+agr]_j T]]]]_{VP} e_i [V_j \dots$

In languages which lack a nominative clitic (like Spanish and Standard Italian), the subject agreement element is used to identify the nom-associated CI:

(19) $[DP_i [CI_i [[[v]agr]_j T]]]_{VP} e_i [V_j \dots$

In other languages, like Standard French, it would seem that either a nominative clitic or the subject-verb agreement element may identify the nom-related CI. See (20a, 20b) below. But upon closer scrutiny, we see that in French (unlike Spanish and Standard Italian), the subject agreement element is not «strong enough» to identify the CI operator on its own. It can only function as an identifier in conjunction with a coindexed lexical DP that is «close enough». See (20c).

- (20) a. Pierre il a vendu sa voiture.
 Pierre nom.cl has sold his car
- b. Pierre a vendu sa voiture.
- c. Pierre sa voiture *(il) l' a vendu.
 Pierre his car (nom.cl) acc.cl has sold

We suggest that the notion of «closeness» invoked above be formalized in the following way:

(21) DP and Agr are close iff they are coindexed and they are contained in the Spec of- or in the head of- adjacent projections at the point of spell-out.

To summarize, I propose to dissociate the CI category from the morphological clitic itself. CI is an abstract operator whose function is to «externalize» an argument of a verb V with respect to the T associated with V. This is achieved by merging a DP with a projection of CI that binds an argument variable of the verb. The function of the clitic contained within T (and, in certain cases, of the verbal agreement affix) is to morphologically identify the abstract CI category. This CI category in conjunction with the Q-operator constitute the linchpin of our analysis of the preverbal subject constraint in Romance interrogatives.

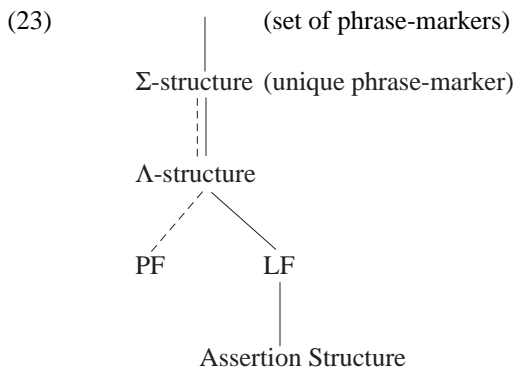
3. The Q-operator and Ph-lexicalization

As mentioned earlier, I assume that an interrogative sentence has a Q feature which encodes the force of the sentence located at the rightmost edge of the CP field. As we shall see in the next section, when the Q feature is coindexed with a certain type of wh-phrase (such as the bare wh-phrases), the Q feature acquires the status of a syntactic operator. I will furthermore assume that a feature that expresses the force of the sentence (such as the Q feature) must be phonologically lexicalized. We define *Ph(onological)-lexicalization* as follows:⁶

- (22) *Ph-lexicalization*: A feature in the head of a functional projection is ph-lexicalized iff the head or the Spec of that projection is associated with appropriate ph-features at the point of spell-out.⁷

We elaborate below on the notion «point of spell-out».

The hypothesis that ph-lexicalization applies at the point of spell out can be formalized within the general architectural framework proposed in Zubizarreta (1998: section 1.3.3). It is argued there that the syntactic derivation proper and the PF interpretation accompany each other for a stretch, namely, the stretch that starts with S-structure (a point in the derivation where the structure involved is a single phrase marker) and ends with L-structure (the last phrase marker in the derivation before it branches). As indicated by the discontinuous line in the diagram below, PF interpretation begins at S-structure and continues after the derivation branches. We will refer to the derivation between S-structure and L-structure as the shallow part of the syntactic derivation and to the structures generated at this point as shallow structures.



6. It is likely that factive predicates, which presuppose the truth of its propositional complement, contain an Ass(ertion) operator in its CP. This operator is lexicalized by the complementizer, which explains why it must be obligatorily present. Cf. *John regrets (*that) Mary is bald*. Complements of propositional attitude verbs lack an Ass operator; therefore, their complementizer may be absent in some languages. Cf. *John thinks (that) Mary is bald*.
7. We shall see later that not only a morpheme but also a feature may fulfill the function of ph-lexicalization. We shall see later that not only a morpheme but also a feature may fulfill the function of ph-lexicalization.

It was argued in Zubizarreta (1998) that the Nuclear Stress Rule, the Focus Prominence Rule, and some prosodically-motivated movements apply in the shallow part of the syntactic derivation. Here we suggest that movement operations triggered solely by the need to ph-lexicalize a category applies in the shallow part of the derivation as well.

There are several ways in which a Q operator may be ph-lexicalized:

- (24) a. by a verbal element:
 (i) via movement of V-to-Q (in direct questions)
 (ii) via merging of Q with matrix V (in indirect questions)
 b. by merging Q with a complementizer
 c. by a wh-phrase in Spec of Q in cases where the Wh-feature is syncretized with Q

As we shall see in section 4, the manner in which the Q-operator is lexicalized will be crucial in understanding the cross-linguistic differences with respect to the preverbal subject constraint in embedded interrogatives. We briefly illustrate below the parameters in (24) and return to it in section 4.

French and English illustrate the parameter in (24a). In direct questions, V moves to Q (via T), where Q is the head of a functional projection at the rightmost edge of the CP:⁸

- (25) a. Qui Pierre a-t-il rencontré hier?
 a'. [Qui [Pierre [a_i-t-il [Q]]] [_{IP} e_i rencontré hier]]
 b. Who did Peter meet yesterday?
 b'. [Who [did_i [Q]]] [_{IP} Peter e_i meet yesterday]]

V-to-Q also applies in yes/no questions:

- (26) a. A-t-il mangé?
 a'. [A_i-t-il [Q]] [_{IP} e_i mangé]
 b. Did he eat?
 b'. [Did_i [Q]] [_{IP} he e_i eat]

In indirect questions, where there is no V-movement to the CP field, we may assume that Q gets lexicalized by merging with matrix V:

- (27) a. Je me demande qui Jean a rencontré hier.
 a'. Je me [Q [demande]] qui Jean a rencontré hier.
 b. I wonder who John met yesterday.
 b'. I [Q [wonder]] who John met yesterday.

In many Northern Italian dialects, V-to-Q also applies in matrix questions, but in embedded questions Q merges with a complementizer at the rightmost edge of

8. In the next section, we discuss the position of the subject in sentences like (25a).

- (31) Me preguntaron que a sus padres_i, qué querían
 to-me asked-3rd-pl that to their parents, what would-like-3rd-pl
 los niños regalarles_i para Navidad.
 the children give-dat.cl for Christmas

This fact may be taken as evidence that syncretism of the Wh and Q features takes place in Spanish. In effect, while the fact that the subordinating conjunction appears to the right of the wh-phrase in Monese can be taken as evidence that the complementizer and the Q feature occupy one and the same position, the fact that the subordinating conjunction appears to the left of the wh-phrase in Spanish may be interpreted as indicating that the Wh and Q features occupy one and the same position. This correlation is sustained by the assumption that the Q feature occupies the rightmost edge of the CP field and that this position is fixed. In other words, features that encode the force of a sentence delimits the (rather weak) frontier between the CP and the IP fields. Therefore, if a feature (such as Wh) syncretizes with Q, or if a complementizer merges with Q, the effects will be seen at the locus of the Q feature (i.e. at the rightmost edge of the CP field). In Spanish indirect interrogatives, the effect is that the wh-phrase appears at the rightmost edge of the CP field and in Monnese indirect interrogatives (as well as in other Northern Italian dialects), the effect is that the subordinating complementizer appears at the rightmost edge of the CP field.

There is a clear generalization that emerges: V-to-C never applies in embedded interrogatives. If this were simply a tendency, we could attribute it to the assumption that «movement» is more costly than «merge». But this explanation cannot account as for why the root/embedded asymmetry is absolute rather than a trend. The analysis proposed above can readily explain that generalization. On the one hand, V-to-C in interrogatives is triggered by the need to ph-lexicalize Q. On the other hand, a movement triggered by the need to ph-lexicalize a category applies at shallow structure. Suppose furthermore that ph-lexicalization, like lexical insertion, is strictly cyclical (where the cyclic nodes are CP and DP). Since the tree is constructed bottom-up and since movement triggered solely by the need to ph-lexicalize a category applies at shallow structure (i.e. at the end of the syntactic derivation), it follows that V-to-Q can only apply on the last cycle (i.e. the root clause).¹⁰

10. As is well-known, V-to-C is blocked in matrix interrogatives when the matrix subject is extracted; cf. **who did leave?* vs. *who left?* A possible explanation is that in such structures one and the same morpheme (namely, Aux) enters into a Spec-head agreement relation with the subject twice: once as the head adjoined to T and another time as the head adjoined to Q (i.e. the wh-phrase moves through the Spec of Q on its way to the Spec of the wh-feature projection). We may speculate that this one-to-many relation between the verb and the subject is not tolerated by the grammar. Or more generally, one and the same lexical head may not enter into multiple Spec-head relations with one and the same argument. This would entail that in languages like English, in which the subject moves to Spec of T, the verb cannot move to T. The correlation seems to be correct: only the feature V moves to T in English, without piedpiping the verbal constituent.

Languages differ as to the default mechanism they turn to in such cases. Some of the Northern Italian dialects resort to the mechanism that they use in embedded questions; i.e. they merge Q

4. The Preverbal Subject in Romance Interrogatives Reconsidered

4.1. The Interaction of Cl and Q Operators: a Minimality Effect

We begin by (re)stating the problem:

- (32) Unlike English, in the Romance languages there are restrictions on the position of the preverbal subject in interrogatives:
- a. Matrix interrogatives do not allow a DP subject to be realized within the IP field: *_{[CP wh Aux _[IP S V]]; *_{[CP wh _[IP S (Aux) V]]}}
 - b. Embedded interrogatives in which the Q morpheme is lexicalized in the CP field (more precisely, at the rightmost edge of the CP field) do not allow a DP subject to be realized within the IP field: *....._{[CP wh Q _[IP S (Aux) V]]}

What is the particularity of the position of the preverbal subject in the IP field in Romance? We suggested in section 2.1 that in Romance the subject is «externalized» from its VP internal position to a position between CP and TP via merge with Cl. Cl is an operator, which binds a variable within the VP. When the Q-feature is coindexed via Spec/head agreement with a [+wh] DP, Q acquires the status of an operator and the trace of the wh-phrase with which it is coindexed acquires the status of a variable:

- (33) a. [_{[wh wh-]_i [Wh_i [e_i [Q_i [....._{[VP e_i....}}

The claim in this paper is that the generalization in (32) is due to the relation between the Cl-operator and the nominative variable it binds, which intervenes between the Q-operator and the variable bound by it. More specifically, we suggest that the generalization in (32) is to be attributed to a minimality effect. To illustrate, consider the French example in (1b) (repeated in (34a) below), which has the structure in (34b). In this structure, the binding relation between the variable e_i and the operator Q is blocked by the presence of the operator Cl, which is closer to e_i than Q.¹¹

with a complementizer giving rise to sentences like *who that left* (see Poletto (in press)). It is not clear what strategy French and English use as a default mechanism in the case of *who left/qui est parti*. We might speculate that they resort to a suprasegmental feature that is spelled out as a rising contour. In other words, what I am suggesting here is that the grammar does not generally rely on intonational properties to lexicalize the Q-operator, but it may do so in certain default cases. See also section 4.1.

11. I have ignored here the so-called stylistic inversion construction; cf. *Qu' a acheté Jean?* 'What bought Jean?'. But the question arises as to whether the postverbal subject occupies the Spec of Cl in such construction (as the analysis put forth by Kayne and Pollock (in press) would suggest), in which case such construction is relevant to our present concerns. Their proposal is that the entire TP is moved into the CP field in this case. If TP moves to a position above the Q-operator (perhaps to the Spec of Q), then there will be no minimality effects in such structures:

(i) [Que_i [WH_i [_{[TP a [e_j acheté e_i]]} [Q [Jean_j [Cl_j [TP]]]]]]]]

- (38) a. A cuánta gente María invitó a la fiesta?
How many people María invited to the party
- b. Combien de gens Marie a invité à la fête?

It is to be noted that such non-bare wh-phrases are precisely the ones that in Romanian require clitic-doubling; cf. Dobrovie-Sorin (1994). This author argues that while bare wh-phrases systematically function as syntactic operators, non-bare wh-phrases do not. The intuition behind Dobrovie-Sorin's proposal may be expressed in the following way:¹³

- (39) A wh-feature in the Spec of DP may or may not percolate to the DP node. If the wh-feature percolates to the DP and if this DP is in the Spec of Q at some point of the derivation, then it will be coindexed with Q. Q will then acquire the status of a syntactic operator and the variable it binds will acquire the status of a variable.

In the sentences in (37) and (38), the wh-feature in the Spec of DP has not percolated to the DP node. Q therefore does not function as an operator and the trace of the wh-phrase does not function as a variable:

- (40) [[_{DP_j} [wh] N] [Q [DP_i [Cl [T [_{VP} e_i [V e_j

Compare with the representation in (41), where the trace of the wh-phrase is bound by Q and therefore has the status of a variable.

- (41) [[_{DP_{j[+wh]}} [wh] N] [Q_j [DP_i [Cl [T [_{VP} e_i [V e_j

Dobrovie-Sorin shows that in the case of «how many N», there are actually two different semantics. If it functions as a syntactic operator (cf. the structure in (41)), it ranges over quantities and if it does not function as a syntactic operator (cf. the structure in (42)), it ranges over objects. This author shows furthermore that *how many N* may have narrow scope with respect to the subject only in the former case (where the trace of the wh-phrase functions as a variable). The prediction is then that in structures where the preverbal subject constraint does not apply, the narrow scope reading should be unavailable. The prediction is borne out: (42b), but not (42a), allows for a distributive reading. *Cada* ('each') being an obligatory distributor, (42a) is uninterpretable.

- (42) a. Cuántos libros María/*cada estudiante leyó ayer?
How many books María/each student read yesterday
- b. Cuántos libros leyó María/cada estudiante ayer?
How many books read María/each student yesterday

13. Dobrovie-Sorin argues that whether or not a wh-specifier percolates its feature to the DP is lexically determined in Romanian (e.g. *which N* vs. *what N*). Our prediction is that while a question with a fronted *which N* in Romanian would not be subject to the preverbal subject constraint, *what N* would be.

The question arises as to how the Q feature is ph-lexicalized in (37) and (38). These sentences have a salient intonational property, namely, there is an intonational hiatus at the right edge of the intonational phrase. More precisely, the pitch falls down to the base line and then rises again, indicating the presence of a prosodic boundary (see Beckman and Pierrehumbert (1986)). We suggest that the prosodic boundary feature is attached to the Q feature and serves the purpose of ph-lexicalizing it.¹⁴

Given what we have said above (cf. (39)), we need to readdress the status of the trace of the wh-phrase in the embedded question in (35). Clearly, the trace of the wh-phrase *que* functions as a variable in embedded questions as much as it does in matrix questions (cf. (1)). We have proposed that the trace of the wh-phrase functions as a variable iff the wh-phrase is associated with Q. When Q defines a projection within the CP field, this association is obtained via Spec/head agreement (as a result of overt movement), but when Q is adjoined to the matrix verb, the association must be obtained in some other fashion. We propose that the mechanism that gives rise to that association is the same one that underlies the association between the NPI and the matrix negative verb in the following type of example:

- (43) Pierre doute que Marie ait vu personne.
 Pierre doubts that Marie has seen anybody

In the example above, the NPI *personne* is licensed by the negative verb *doute* (= 'not sure'). In other words, *personne* and *douter* are associated in some way. We may assume that this is done by covert neg-feature movement. Let us assume that the wh-phrase and the interrogative verb Q+*demander* are associated in the same fashion (via covert wh-feature movement). The resulting representation would then be as in (44) (rather than as in (35b)). In this representation, the lexical interrogative verb functions as the operator that binds the wh-variable (i.e. the trace of *que*).

- (44) Je me [Q [demande]]_i [que_i [Wh_i [Jean_j [Cl_j [_{TP} a [_{VP} e_j [acheté e_i hier]]]]]]]

The question that arises with respect to such a representation is why it does not give rise to a minimality effect. We suggest the following answer. In (44), the operator Cl does not intervene in the relation between the operator *demander* and *e_j* because Cl is a functional category and *demander* is a lexical category. In other

14. As is wellknown, fronted wh-adverbs do not activate the «preverbal subject constraint» in Romance interrogatives:

- (i) Cuando Pedro compró este libro?
 When Pedro bought this book
 (ii) Quand Pierre a acheté ce livre?

Given the role of the Cl operator as an argument-externalizer, a plausible assumption is that Cl does not «see» adverbial positions. Another alternative (proposed in Zubizarreta (1998)) is that wh-adverbs may be directly generated in the CP field.

words, we suggest that minimality should be relativized not only with respect to the distinction between phrases and heads (cf. Rizzi (1990)), but also with respect to types of heads (lexical vs. functional).

4.2. Getting around Minimality: Projecting Cl within the CP Field

Given the minimality effect created by the projection between CP and TP of a nom-associated Cl in interrogative structures like (34), the question arises whether a nom-associated Cl may ever coexist with a Q-projection. The answer is that they can coexist but iff the nom-associated Cl is projected higher than the Q-projection. According to Poletto (in press), in French and in many Northern Italian dialects, the subject may be realized within the CP field, as in the examples below:

- (45) A qui Pierre a-t-il parlé?
 to whom Pierre has-he (nom.cl) spoken
 'To whom has Pierre spoken?' (French)
- (46) I me ga domandà Gianni quando che el vien
 nom.cl me have asked John when that nom.cl comes
 'They asked me when Gianni is coming.' (a Northern Italian variety)
- (47) A venta che Majo ch' a mangia pi' tant
 cl need that Majo that nom.cl eat more
 'Majo must eat more.' (Piedmontese)

The possibility of realizing the subject within the CP field might actually be an option more widely available within the Romance family. This option is made available by the fact that the subject is externalized via Cl in these languages. While Cl is generally projected at the left-edge of the IP field (i.e. immediately above TP), the grammar allows for the option of projecting Cl within the CP field in order to avoid a minimality violation, namely above the Q operator and below the Wh-projection in French but above it in the Northern Italian dialects. Thus, the French example in (45) will be associated with the structure below (before V-to-Q movement applies in shallow structure):

- (48) [A qui_i [Wh_j [Pierre_i [Cl_i [e_j [Q_j [il_i a [e_i [parlé e_j]]]]]]]]]]
-

In (48), Cl is above Q; therefore Cl does not interfere in the binding relation between Q and e_j. The question that arises is why Q does not interfere in the binding relation between Cl and the variable e_i. I suggest that this is due to the so-called anti-agreement effect discussed by Ouhalla (1993) and references cited therein. We elaborate below.

A variable bound by Cl has the status of a pronominal variable. This is due to the fact that the clitic or «strong» agreement element that identifies Cl is [+pronominal]. There is a parametric variation among languages:

The other question is why the nominative clitic is an enclitic in (45). Recall that the nominative clitic originates above V (see section 2 above). Therefore it cannot adjoin to V; instead, it left-adjoins to T (after V left-adjoins to T), giving rise to the following structure:

(52) $[_T \text{ nom cl } [_T [V] T]]$

When V-to-Q applies in shallow structure for the purpose of ph-lexicalization, it only moves the category with ph-content, namely V (i.e. T has no ph-content):


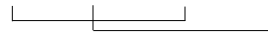

(53) $[[_V_i [Q]] [_T \text{ nom cl } [_T [V e_i] T]]]$

The nominative clitic will then attach at PF to the verbal category to its left. Our claim that the nominative clitic may attach to a word to its left by a phonological, sandhi-type rule, receives some support from the Northern Italian dialects where the nominative clitic may optionally attach to a complementizer to its left (cf. Poletto (in press)):¹⁶

- (54) a. Ara che el vien. (Veneto)
 Look that nom.cl comes
 b. Ara ch'el vien

4.3. Similarities with Chinese with respect to Minimality

It is timely to mention the case of the indefinite wh in Mandarin Chinese, discussed by Li (1992). Li shows that wh-words in Chinese can have a non-interrogative indefinite interpretation as well as an interrogative one, depending on the semantic properties of its binder. The relation between a non-interrogative wh and its licensor is a binder-variable relation, like the one that exists between an interrogative wh and a question operator. And these binder-variable relations are subject to Minimality, as illustrated by the configurations below:

- (55) a. +QP -QP wh1 wh2

 b. *+QP -QP wh1 wh2

 c. *+QP -QP wh1 wh2


16. Sportiche (1998) claims that the nominative enclitic has a tighter morpho-phonological connection to the verb than a nominative proclitic and therefore proposes that enclitics be lexically generated with the verb. I would suggest that this is a more general morphological property of words: morphemes to the right of the stem are closer to the stem than the morphemes to the left of the stem. This is true not only of clitics but also of affixes. Cf. the possibility of coordinating prefixes but not suffixes; e.g. *pre- and post-position*.

syntactic predicate; a pronominal or lexical DP merges with Cl and functions as the subject of the predicate. Cl is thus formally comparable in status to a relative operator. Cl is morphologically identified by a pronominal clitic or strong agr contained in T. Languages with clitics and strong agr, like many of the Romance languages, thus have a mechanism other than movement to «externalize» an argument, in particular the subject. The proposal then is that preverbal (nonfocused) subjects in Romance are in the Spec of a Cl-operator and this Cl creates minimality effects when a question is formed by moving a wh-phrase above it. More precisely, a wh-phrase moves to the Spec of a Wh-projection through the Spec of a Q-projection (located at the rightmost edge of the CP field). This creates a coindexing relation between the wh-phrase and Q, which then binds the trace of the wh-phrase as well, giving rise to a Q-operator/variable relation. But the Cl operator is closer to the wh-variable than the Q-operator, a minimality violation thus arise.¹⁷

The absence of the preverbal subject constraint in interrogatives with fronted nonbare wh-phrases was attributed, following Dobrovie-Sorin (1994), to the fact that the wh-feature in the specifier of such phrases may fail to percolate to the DP, in which case there is no coindexing of a wh category with Q. In such structures, Q does not function as a syntactic operator and no minimality ensues.

While Cl is generally mapped in the higher part of the middle field, namely between the Q projection and TP, in certain cases Cl may project above the Q projection, thus avoiding a minimality effect. This is the case of the so-called complex inversion construction in French and of a similar construction in the Northern Italian dialects.

It was suggested that the differences among the Romance languages with respect to the preverbal subject constraint in embedded interrogatives is to be attributed to parametric differences in the ph-lexicalization of the Q-feature that expresses the force of the sentence.

Appendix: On Acc Clitic Doubling in Castilian Spanish¹⁸

In Castilian Spanish an accusative clitic can double a wh-phrase only in the presence of an active negation or in a modal context:

- (60) a. *A quién no lo aguanta ni el más paciente*
 Whom (obj) not acc.cl stand-3rd-sing even the most patient
de los mortales/ni su madre. (Active Neg)
 of the mortals/even his mother (subj)

17. An exception is Brazilian Portuguese. In effect, the preverbal subject constraint is absent in Brazilian interrogatives. This is as expected. Brazilian Portuguese lacks Cl because it has no way of morphologically identifying it: it has a weakened subject agreement system and has lost (or is in the process of losing) the 3rd person accusative clitic. See Marcello Modesto (1999) and references cited therein.

18. I would like to thank Marina Fernandez Lagunilla for providing judgements on the data reported in this section.

- b. (?) *A quién* no *lo* aguanta nadie? (Active Neg)
Whom (obj) not acc.cl stand-3rd-sing nobody (subj)
- c. **A quién* no *lo* aguanta María (subj) (Non-active Neg)
- (61) a. *A quién* (piensas que) *lo* debería castigar
Whom (obj) (think-2nd-sing that) acc.cl should-3rd-sing punish
María/ su madre? (modal verb / «deontic»)
María/ his mother (subj)
- b. *A quién* dudas de que *lo* castigue
Whom (obj) doubt-2nd-sing of that acc.cl punish-3rd-sing-subjunc
María/ su madre? (present subjunctive)
María/ his mother (subj)
- c. *A quién* (piensas que) *lo* castigará
Whom (obj) (think-2nd-sing that) acc.cl punish-2nd-sing-fut
María/ su madre? (future indicative)
María/ his mother (subj)
- (62) a. **A quién* (piensas que) *lo* castigó
Whom (obj) (think-2nd-sing that) acc.cl punish-3rd-sing-past
María/ su madre? (past indicative)
María/ his mother (subj)
- b. **A quién* (piensas que) *lo* castiga
Whom (obj) (think-2nd-sing that) acc.cl punish-3rd-sing-pres
María/ su madre? (present indicative)
María/ his mother (subj)

Some remarks about the data are in order:

1. There is no specificity effect: the wh-phrase that is being doubled is not interpreted as being discourse-linked.
2. The wh-phrase (*a quién*) can bind the postverbal subject without giving rise to WCO effects. This is as expected since the wh-phrase originates in the Spec of Cl. See section 2.
3. The accusative clitic being +definite, it renders the question +definite. In other words, existence of an individual corresponding to the object is presupposed by both speaker and hearer. This is shown by the fact that «nobody» is not a possible answer for any of the questions in (60) and (61).

I would like to propose that in order to understand the phenomenon described above, we must appeal to the level of Assertion Structure in the sense of Zubizarreta (1998). The Ass Structure is a post-LF level where the partitioning of the sentence in terms of focus and presupposition is articulated in terms of ordered proposi-

tions.¹⁹ The (shared) presupposition is an assertion that introduces an existentially bound variable. The focus of the sentence provides a value for the variable introduced in the presupposition. The former therefore precedes the latter in the Ass Structure. To illustrate, consider the example in (63a), with the F(ocus)-marking as indicated in (63b). This sentence will be associated with the Ass Structure given in (63c).

- (63) a. María invited Juan.
 b. María invited [_F Juan]
 c. Ass Structure:
 Ass (1): There is an x such that María invited x
 Ass (2): The x such that María invited x = Juan

It was noted in Zubizarreta (1998) that the relation between the definite description in Ass (2) and the existentially-bound variable in Ass (1) is formally comparable to the relation between an E-type pronoun and an indefinite in a previous sentence, as in the example below.

- (64) A sailor came in. *He* was wearing a red shirt.
 [he = the sailor that came in]

Capitalizing on this intuition, we suggest that the accusative clitic in (60) and (61) has the status of an E-type pronoun. We develop it below.

Suppose that negation and modals introduce an independent assertion. Consider the examples in (60a) and (61a). The presence of negation and the modal give rise to Ass (1) in (65) and (66) respectively. Ass (2) assigns a value (i.e. a constant) to the variable introduced in Ass (1). We use a capital letter to indicate a constant, where the constant corresponds to the possible answers of the question. Our claim is that the accusative clitic corresponds to the definite description in Ass (2).

- (65) Ass (1): There is NO x (not even the most patient of mortals) such that x stands some y

Ass (2): The y such that there is no x...(x stands y) = A
 _____ lo _____

- (66) Ass (1): SHOULD (María punish some y)

Ass (2): The y such that (should (María punish y) = A
 _____ lo _____

19. I have also argued in the same work that the notions topic/comment should be articulated at the Ass Structure level in terms of the subject/predicate relation.

We then reach the following conclusion for Castilian Spanish:

- (67) In Castilian Spanish, an accusative clitic may be QU-bound iff the clitic has the status of an E-type pronoun (in other words, iff the accusative clitic corresponds to a definite description in the Ass Structure of the sentence).

References

- Alexiadou, A.; Anagnostopoulou, E. (1998). «Parametrizing AGR: Word Order, V-Movement, and EPP-Checking». *Natural Language and Linguistic Theory* 16: 491-539.
- Aoun, J.; Benmamoun, E. (1998). «Minimality, reconstruction, and PF movement». *Linguistic Inquiry* 20: 141-172.
- Barbosa, P. (1995). *Null Subjects*. Department of Linguistics and Philosophy. Cambridge: MIT. PhD dissertation.
- Benincà, P.; Poletto, C. (1997). «A case of *do* support in Romance». University of Padua. Unpublished manuscript.
- Cecchetto, C. (1995). «Reconstruction in clitic-left dislocation». Department of Cognitive Science, Istituto scientifico San Raffaele, Milan. Manuscript.
- Chomsky, N. (1993). «A minimalist program for linguistic theory». In Hale, K.; Keyser, S. J. (eds.). *The View from Building 20*. Cambridge, Mass.: MIT Press.
- (1994). «Bare phrase structure». *MIT Occasional Papers in Linguistics* 5. Department of Linguistics and Philosophy. Cambridge, Mass.: MIT [Also in Webelhuth, G. (ed.). *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell, 1995.]
- Dobrovie-Sorin, C. (1994). *The Syntax of Romanian*. Mouton de Gruyter.
- Fernández-Soriano, O. (1989). *Rección y ligamento en español: Aspectos del parámetro de sujeto nulo*. Universidad Autónoma de Madrid. Doctoral dissertation.
- Franco, J. (1993). *On object agreement in Spanish*. Department of Spanish and Portuguese. University of Southern California, Los Angeles. Doctoral dissertation.
- Guasti, M. T. (1996). «On the controversial status of Romance interrogatives». *Probus* 8: 161-180.
- Iatridou, S. (1990). «Clitics and island effects». Cambridge, Mass.: MIT. Unpublished manuscript.
- Kayne, R. (1991). «Romance clitics, verb movement, and PRO». *Linguistics Inquiry* 22: 647-686.
- (1994). *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Koopman, H. (1997). «Unifying predicate cleft constructions». Paper presented at *BLS* special session.
- Li, A. (1992). «Indefinite wh in Mandarin Chinese». *Journal of East Asian Linguistics* 1: 125-155.
- Modesto, M. (1999). «Null subjects without rich agreement». To appear in Kato, M.; Negrão, E. (eds.). *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt and Madrid: Vervuert-Iberoamerican.
- Munaro, N. (1997). *Proprieta strutturali e distribuzionali dei sintagmi interrogativi in alcuni dialetti italiani settentrionali*. Dipartimento di Linguistica, University of Padua. Unpublished dissertation

- Ouhalla, J. (1993). «Subject-extraction, negation and the anti-agreement effect». *Natural Language and Linguistic Theory* 11.
- Poletto, C. (in press). *The Higher Functional Field in the Northern Italian Dialects*. Oxford Studies in Comparative Syntax. Oxford University Press.
- Pollock, J.Y.; Murano, N.; Poletto, C. (1998). «Eppur si muove! On comparing French, Portuguese and Bellunese Wh-movement». CNRS and University of Padua. Unpublished manuscript.
- Rizzi, L.; Roberts, I. (1989). «Complex Inversion in French». *Probus* 1: 1-30.
- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- (1991). «Residual verb second and the *Wh*-Criterion». *Faculté des Lettres, Université de Genève. Technical Reports in Formal and Computational Linguistics* 2.
- Speas, M. (1994). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Sportiche, D. (1998). *Partitions and Atoms of Clause Structure*. Routledge.
- Suñer, M. (1988). «The role of agreement in clitic-doubled constructions». *Natural Language and Linguistic Theory* 6: 391-434.
- (1992). «Subject clitics in the Northern Italian vernaculars and the Matching Hypothesis». *Natural Language and Linguistic Theory* 10: 596-640.
- (1994). «V-movement and the licensing of argumental wh-phrases in Spanish». *Natural Language and Linguistic Theory* 12: 335-372.
- Zubizarreta, M.L. (1993). «The grammatical representation of topic and focus: Implications for the structure of the clause». *Cuadernos de Lingüística* 2. Instituto Universitario Ortega y Gasset.
- (1998). *Prosody, Focus, and Word Order*. Cambridge, Mass.: MIT Press.