

The when and where-to of subject NP inversion: implications for an HPSG grammar of Spanish

Vanessa Metcalf

June 5, 2003

Subject NP inversion in Spanish

The term *subject NP inversion* is used to refer to cases where the subject of a clause appears somewhere to the right of the matrix verb. In Spanish, subject NP inversion is optional in a variety of clause types.

- (1) a. Contestó la pregunta **Juan**.
answered the question Juan
'Juan answered the question.'
- b. **Juan** contestó la pregunta.
Juan answered the question (Torrego 1984)
'Juan answered the question.'

Obligatory subject NP inversion

Obligatory subject NP inversion: The subject cannot appear to the left of the finite verb. According to Torrego (1984), this is the case in embedded and non-embedded *wh*-interrogative clauses with preposed *wh*-phrases.

Obligatory inversion in an embedded interrogative

- (2) a. No sabía [qué querían esos dos]
not knew what wanted those two
'I didn't know what those two wanted'
- b. * No sabía [qué **esos dos** querían]
not knew what those two wanted (Torrego 1984)

Obligatory inversion in a non-embedded interrogative

- (3) a. ¿Qué querían **esos dos**?
what wanted those two
'What did those two want?'
- b. * ¿Qué **esos dos** querían?
what those two wanted (Torrego 1984)

Overview

An adequate account of subject NP inversion must:

- be able to characterize the syntactic environment in which inversion is obligatory, and enforce it (the when)
- allow for the word order possibilities associated with inversion, while ruling out impossibilities (the where-to)

Last year in Colloquium Fest, I presented work towards such an account, which relied on linearization in accounting for word order. Unfinished business included two issues I will address today:

- Some data pertaining to obligatoriness was not explained.
- The possibility of accounting for the word order facts using argument composition had not been ruled out.

Defining the relationship between obligatory inversion and extraction: Part 1

Subject NP inversion is not obligatory in *wh*-relative clauses, which superficially look much like *wh*-interrogatives.

- (4) el libro [que **Juan** escribió]
the book that Juan wrote
'the book that Juan wrote'

Thus, the presence of a clause-initial interrogative *wh*-phrase triggers obligatory inversion, but the presence of a clause-initial relative *wh*-phrase does not.

The visibility of QUE and REL

The distinction between interrogative and relative *wh*-phrases is encoded in HPSG via the *nonlocal* features QUE and REL, respectively.

A finite verb whose maximal projection is a head-filler construction has access to information about the QUE and REL values of the filler via the *head* path. That is, it is possible to formulate a constraint relating HEAD feature specifications of a phrase to NONLOCAL feature specifications of its non-head daughter.

A finite verb heading an embedded clause along an extraction pathway does not have access to this information, since:

- REL and QUE are *nonlocal* features
- SLASH lists contain *local* objects

The plain English version of the last slide

HPSG feature structures organize information such that:

Information about whether a filler is an interrogative *wh*-phrase or a relative *wh*-phrase is available at the level where it is realized.

It is possible to tie *head* feature specifications of the lexical head of that clause to this information, since the *head* feature is percolated down to the finite verb.

Further down a filler-gap pathway, however, lexical heads cannot “see” whether the filler higher up is an interrogative *wh*-phrase or a relative *wh*-phrase.

Prediction of the theory w.r.t obligatory inversion and an extraction pathway

Researchers have noted a number of syntactic phenomena that mark each step along an extraction pathway (Hukari and Levine 1995), and are always on the lookout for more, but . . .

The assumptions outlined above predict that a syntactic phenomenon involving finite verbs cannot both:

- be associated with every step along an arbitrarily long extraction pathway
- depend upon the presence of a filler with a non-empty QUE value at the top of that pathway

Support for (or at least no complaints about) the HPSG treatment of *nonlocal* features

Evidence pertaining to obligatory inversion is consistent with this prediction: the subject of a finite verb in an embedded clause along an extraction pathway is not obligatorily inverted.

- (5) ¿Qué_i pensaba **tu** **padre** [que **Alberto** le daría _{-i} a
what_i thought your father that Alberto to-him would-give _{-i} to
Juan]?
Juan

'What did your father think that Alberto would give to Juan?'

In (5) the subject of the matrix clause is inverted, but the subject of the embedded clause, from which *qué* is extracted, need not be.

Is obligatory inversion extraction-insensitive?

All the evidence discussed thus far points to an account of obligatory inversion in which it is not necessary to make reference to extraction (i.e. non-empty SLASH values or head-filler constructions) at all.

Assuming that subject inversion depends on the familiar boolean verbal *head* feature INV, a finite clause with an interrogative *wh*-phrase non-head daughter can be required to exhibit subject NP inversion.

$$\left[\begin{array}{l} \text{SS|LOC|CAT} \left[\begin{array}{l} \text{HEAD|VFORM } \mathit{fin} \\ \text{VAL} \left[\begin{array}{l} \text{SUBJ} \quad \langle \quad \rangle \\ \text{COMPS} \langle \quad \rangle \end{array} \right] \end{array} \right] \\ \text{NON_HD_DTR|SS|NONLOC|QUE } \mathit{ne_list} \end{array} \right] \rightarrow [\text{SS|LOC|CAT|HEAD|INV } +]$$

Figure 1: Enforcing inversion in *wh*-interrogatives without reference to extraction

Defining the relationship between obligatory inversion and extraction: Part 2

However, the generalization that all clause-initial interrogative *wh*-phrases condition obligatory inversion is incorrect, since both *por qué* and *cómo* permit non-inverted structures.

- (6) a. ¿Por qué dijo **Juan** que le dado el libro a Alberto?
 why said Juan that CL(to-him) had given the book to Alberto
 ‘Why did Juan **say** that he had given the book to Alberto?’
 ‘Why did Juan say that he had **given** the book to Alberto?’
- b. ¿Por qué **Juan** dijo que le había dado el libro a Alberto?
 why Juan said that CL(to-him) had given the book to Alberto
 ‘Why did Juan **say** that he had given the book to Alberto?’

Uninverted *wh*-interrogatives have a more restricted interpretation than their inverted brethren

The reading of (6a) which involves questioning why something was given necessarily involves extraction, in order to associate *por qué* with the embedded clause. This reading is only available in the context of subject NP inversion.

A straightforward account of this fact is possible given the following assumptions:

- *Por qué* can combine with an ungapped clause to form an interrogative clause (which is not an instance of a head-filler construction).
- Obligatory subject NP inversion is an extraction-sensitive phenomenon in the sense that only *wh*-interrogatives which are head-filler constructions are obligatorily inverted.

Enforcing obligatory inversion: revised constraint

$$\left[\begin{array}{l} hd_fill_phr \\ NON_HD_DTR|SS|NONLOC|QUE \ ne_list \end{array} \right] \rightarrow [SS|LOC|CAT|HEAD|INV \ +]$$

Figure 2: Obligatory inversion as an effect of the syntactic environment

This revised constraint states that a head-filler construction whose filler daughter has a non-empty `QUE` value must have the *head* specification `[INV +]`.

Word order in inverted clauses

The subject NP in an inverted clause may:

- be interspersed with the complements of the matrix verb (7a and 7b)
- appear within a complement VP (7c)

- (7) a. Quiere [comprar el libro]_{VP} **Juan**?
wants buy the book Juan
- b. Quiere **Juan** [comprar el libro]_{VP}?
wants Juan buy the book
- c. Quiere [comprar **Juan** el libro]_{VP}?
wants buy Juan the book
- ‘Juan wants to buy the book.’

The problem posed by inverted subjects

Examples like (7c), in which the subject is realized within a complement VP, pose a problem for an HPSG account making the following assumptions:

1. A finite verb realizes a subject.
2. An embedded verb and its (non-subject) arguments form a constituent.
3. Phonological form is derived directly from constituent structure.

An account of this data needs to dispense with one of these assumptions, and provide an alternative.

Possible explanations for word order

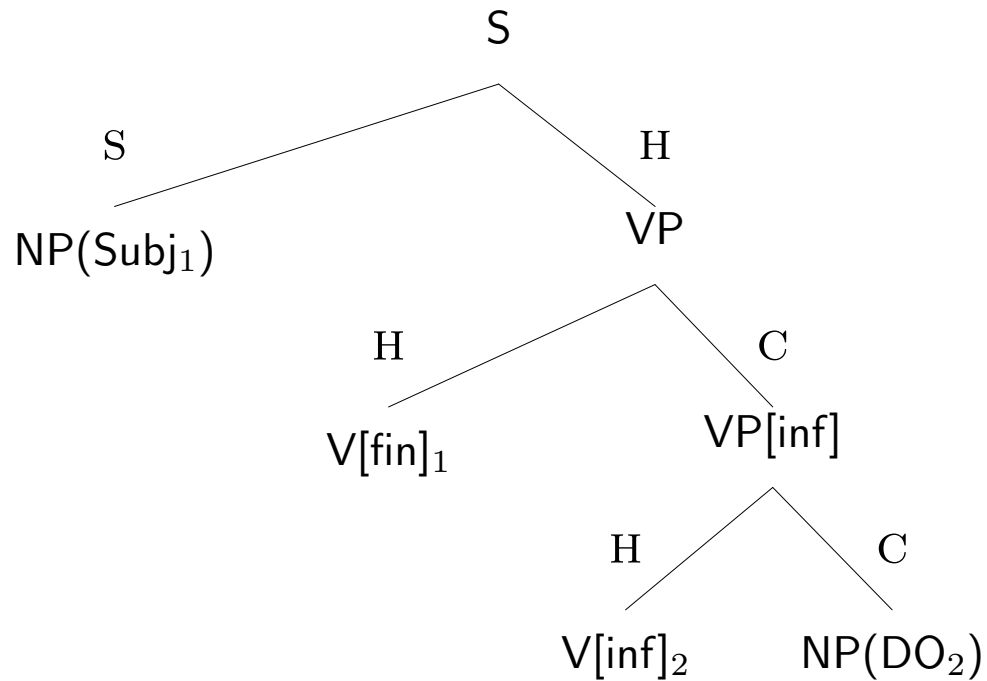
Given a phonological form corresponding to the string

(8) V[fin]₁ V[inf]₂ NP(Subj₁) NP(DO₂)
quiere comprar Juan el libro

where NP(Subj₁) is taken to be the subject of the finite verb V[fin]₁, and NP(DO₂) is taken to be a complement of the infinitive verb V[inf]₂, I see three possibilities.

Possible structure 1 for $V[\text{fin}]_1 V[\text{inf}]_2 \text{NP}(\text{Subj}_1) \text{NP}(\text{DO}_2)$

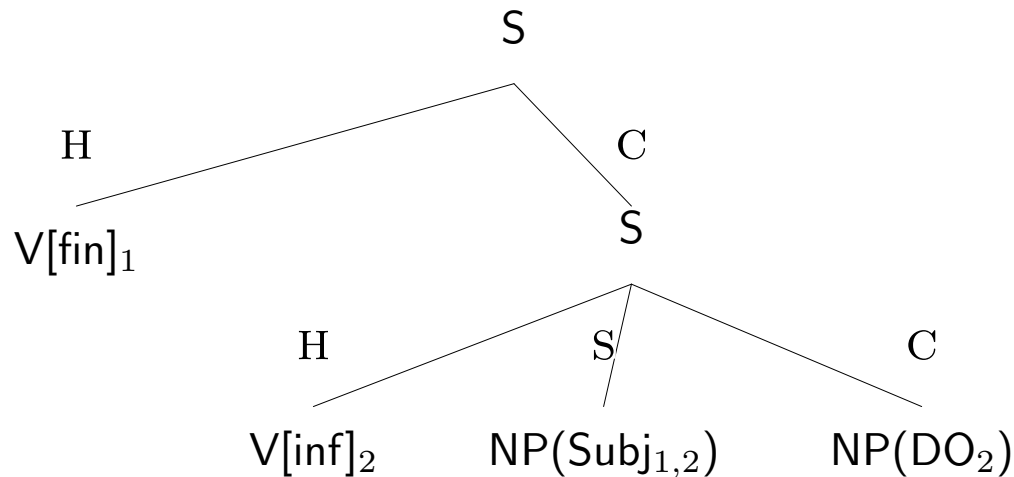
Maintain the traditional constituent structure of a clause, but map it to phonological form using the mechanisms of a linearization theory



Possible structure 2a for V[fin]₁ V[inf]₂ NP(Subj₁) NP(DO₂)

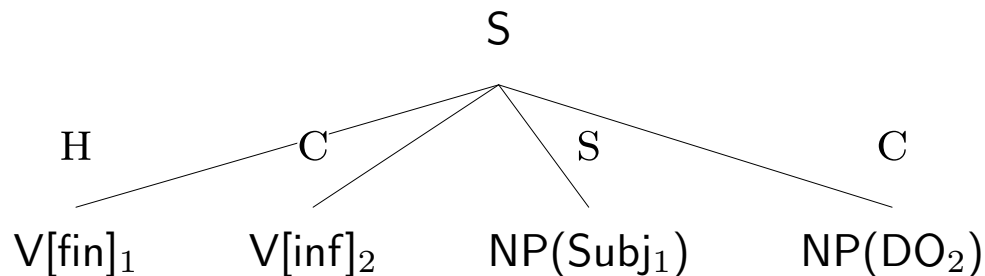
Maintain a direct mapping of constituent structure to phonological form, but posit a different constituent structure, in which the VP[inf]₂, NP(Subj₁), and NP(DO₂) either:

a) form a constituent which is a dependent of the VP[fin]



**Possible structure 2b for
V[fin]₁ V[inf]₂ NP(Subj₁) NP(DO₂)**

b) do not form a constituent, and are individual dependents of the VP[fin]



This structure can be licensed based on *argument composition* or *argument inheritance* (Hinrichs and Nakazawa 1989).

In argument composition, a verb that normally selects for a VP complement may select for a lexical verb and its complements. This variation is explained in terms of two lexical entries, one derived from the other via a lexical rule.

Example of an argument composition LR

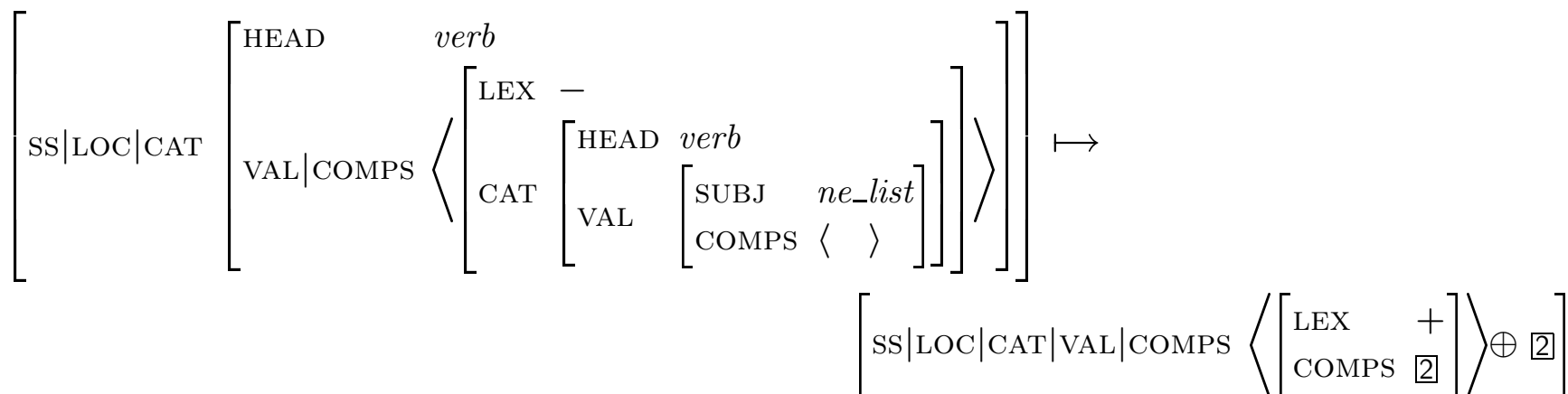
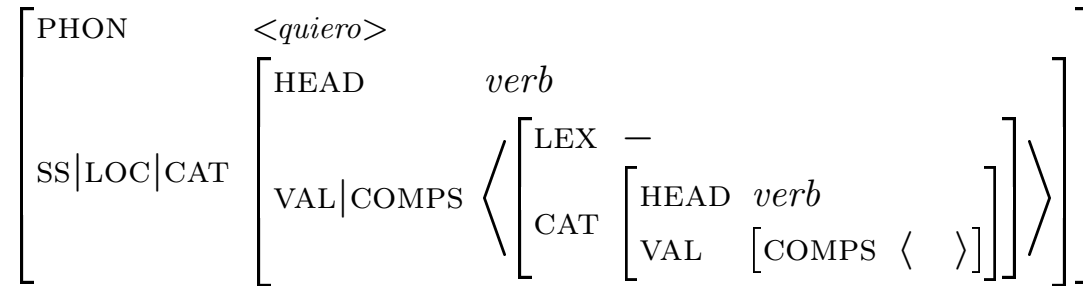


Figure 3: Argument composition lexical rule

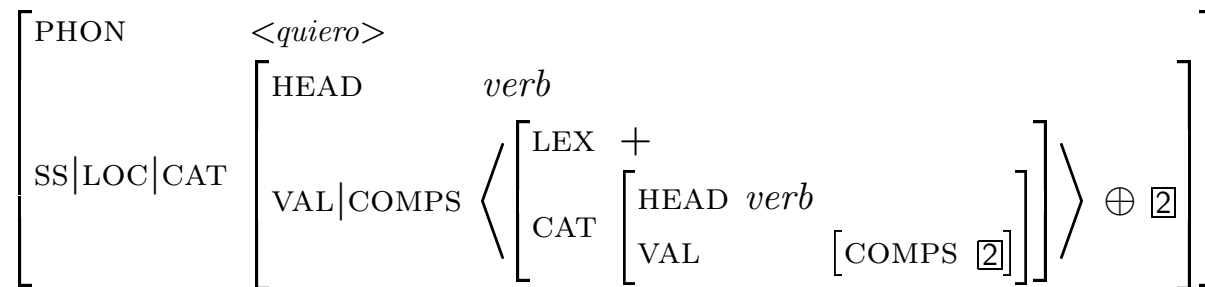
Syntactic tests suggested by Monachesi (1998), involving clitic attachment and a passive construction, are consistent with the existence of such a lexical rule in Spanish.

The two kinds of *quiero*

Selecting a VP argument



Selecting a verbal head + argument composition



Background: Pronominal cliticization in Spanish

I will provide some background about Spanish clitics, in order to later explain data pertaining to argument composition which involves cliticization.

- (9) a. Juan vio **a Alberto**.
Juan saw Alberto.
'Juan saw Alberto.'
- b. Juan **lo** vio.
Juan CL-him saw
'Juan saw him.'
- c. Quiero comprar **el libro**.
want buy the book
'I want to buy the book.'
- d. Quiero comprar**lo**.
want buy-CL(it)
'I want to buy it.'

Direct objects (but not subjects) of a verb are optionally realized as clitics. Clitics precede finite verbs (9b), and follow infinitive verbs (9d). Other verbal forms are not relevant here.

Background: Clitic climbing

In some cases, clitics may occur with either a higher verb, or an embedded verb. Whether or not such variation is possible depends on the higher verb, which is underlined in the examples below.

- (10) a. **Lo** quiero comprar.
CL(it) want buy
'I want to buy it.'
- b. Quiero comprar**lo**.
want buy-CL(it)
'I want to buy it.'

Most auxiliary verbs show this pattern of variation, in addition to a class of non-auxiliary verbs that can apparently vary from speaker to speaker. Such verbs are referred to as *restructuring verbs*.

Background: Passive *se*

In addition to the pronominal clitics discussed above, the clitic *se* can function as a passive marker on verbs with active morphology.

- (11) a. El gobierno ha decretado nuevas leyes.
the government has(3sg) enacted new laws
'The government has enacted new laws'
- b. **Se** han decretado nuevas leyes.
CL-passive have(3pl) enacted new laws
'New laws have been enacted'

These examples are provided by M. Aránzazu Martín-Lozano (ms.).

Background: Passive *se*

This construction is called passive rather than impersonal when the verb agrees in person and number with the NP that would ordinarily be the direct object, and that NP is considered to be the subject.

Thus, in (11a) the subject is *el gobierno*, and the verb shows third person singular agreement; in (11b) the clitic *se* is attached to the verb and the verb shows third person plural agreement with *nuevas leyes*.

Background: Passive *se*

Further evidence that the NP *nuevas leyes* should be treated as a subject is provided by the fact that it cannot be realized as a clitic; cliticization of subjects in Spanish is ungrammatical.

- (12) * **Se** las han decretado.
CL-passive CL-them have(3pl) enacted
'They (the laws) have been enacted.'

Restructuring verbs can occur with passive *se*, as shown in example (13), where *los libros* is assumed to be the subject of *quieren...*

- (13) **Se** quieren quemar los libros.
CL-passive want(3pl) burn the books
'The books are wanted to be burned'

The interaction of clitic attachment and passive *se*

... But only the highest verb can occur with clitics in the presence of passive *se* (14a); attachment to the embedded verb is ungrammatical (14b), though fine in an active construction (14c).

- (14) a. **Se** **le** quieren dar los libros.
CL-passive CL(to-him) want(3pl) give the books
'The books are wanted to be given to him'
- b. * **Se** quieren darle los libros.
CL want(3pl) give-CL(to-him) the books
- c. Quiero darle los libros.
want give-CL(to-him) the books
'I want to give him the books'

When clitics climb

Most analyses of passive constructions hold that an NP is only eligible to be the subject of a passivized verb when it could have been the complement of some active incarnation of that verb.

The grammaticality of (14a) supports the notion that there is a lexical entry for *quieren* in which *los libros* is on its COMPS list. A passive lexical rule would take that lexical entry as input in order to license (14a).

Assume this lexical entry is derived via argument composition.

Since clitic attachment to *quieren* co-occurs with *los libros* as a subject of *quieren*, the case in which we assume argument composition must have applied, clitic climbing can be used as a sort of flag for argument composition.

When clitics don't climb

On the other hand, clitic attachment to the embedded infinitive verb *regalar* cannot co-occur with *los libros* as a subject of *quieren*.

This suggests that argument composition, a necessary precondition for passivization in this case, has the consequence that the verbal complement cannot realize clitics.

In cases where an embedded verb does realize clitics, as in (14c), it follows that argument composition has not occurred, and so the embedded verb must also realize its overt complements.

Thus, clitic attachment to an embedded verb followed by non-subject arguments points to a VP constituent headed by that embedded verb.

Structure and clitic attachment

In summary, the preceding data supports the association of clitic placement with the structures shown in figure (4).

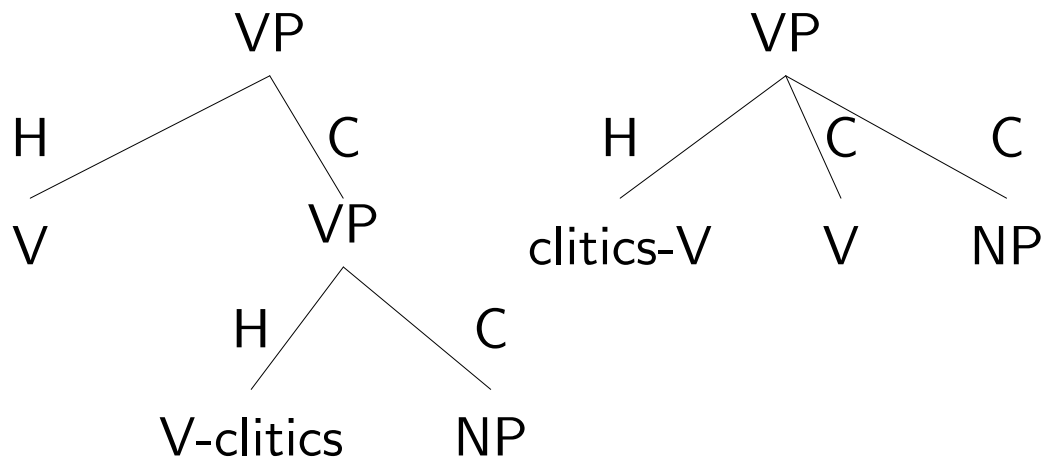


Figure 4: The interaction of cliticization and argument composition

Predictions for subject inversion word order

A flat-structure account mapping constituent structure directly to phonological form can easily account for a subject being interspersed with the complements of the verb that selects it, but not occurring within one of those complements.

The occurrence of a subject NP between an embedded verb and one of its arguments is predicted to be:

- grammatical where argument composition has occurred (signaled by clitic attachment to the higher verb)
- ungrammatical where argument composition has not occurred (signaled by clitic attachment to the embedded verb)

Argument composition is not an adequate explanatory device for inverted subject word orders

In fact, as shown in example (15), both possibilities are fine.

- (15) a. **Se** **lo** quiere regalar **Juan** a Pedro?
CL(to-him) CL(it) want(3sg) buy Juan to Pedro
'Does Juan want to buy it for Pedro?'
- b. Quiere [regalarse**lo** **Juan** a Pedro]_{VP}?
want(3sg) buy-CL(to-him)-CL(it) Juan to Pedro
'Does Juan want to buy it for Pedro?'

Even in the case that proponents of argument composition would assume involves a VP constituent (15b), the inverted subject can appear within the VP. Argument composition cannot explain this word order possibility.

Summary and Outlook

1. A straightforward account of obligatory subject NP inversion is possible if it is assumed that not all *wh*-interrogative clauses are unbounded dependency construction.
2. The word orders associated with subject NP inversion in general cannot be explained by traditional constituency-based accounts, nor can they be explained by positing the kind of flat structure provided by argument composition.
3. This leaves two remaining possibilities in accounting for word order facts:
 - delinking constituent structure from phonological form, mapping one to the other via the mechanisms provided by linearization
 - or, an alternative constituency-based account needs to be developed, in which an embedded verb may realize not only its complements, but also the subject of the matrix verb

References

- Hinrichs, E. W. and T. Nakazawa (1989). Subcategorization and VP structure in German. In *Aspects of German VP Structure*, Sfs-Report-01-93. Ebershard Karls Universitat Tubingen.
- Hukari, T. and R. D. Levine (1995). Adjunct extraction. *Journal of Linguistics* 31, 195–226.
- Monachesi, P. (1998). Decomposing Italian clitics. In S. Balari and L. Dini (Eds.), *Romance in HPSG*, pp. 305–357. Stanford: CSLI publications.
- Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1), 103–129.