Spanish postverbal subject NPs: Towards an HPSG account of obligatoriness and word order *

Vanessa Metcalf
The Ohio State University
vmetcalf@ling.ohio-state.edu

April 19, 2004

Abstract

Postverbal subject NPs, the occurrence of a subject NP to the right of the finite verb which selects it, is usually optional in Spanish clauses, but in some cases obligatory. In this paper, I characterize both the syntactic environment in which postverbal subjects are obligatory, and describe the word order possibilities associated with postverbal subject clauses in general, providing an account of both obligatoriness and word order within the HPSG framework. With respect to obligatory postverbal subjects, it is shown that a straightforward characterization of the syntactic environment conditioning Spanish obligatory postverbal subject NPs is possible if it is assumed that not all *wh*-interrogative clauses are unbounded dependency constructions. Furthermore, it is shown that predictions which follow from the HPSG treatment of unbounded dependencies are consistent with data from both Spanish and French. With respect to word order, Spanish postverbal subject clauses pose a problem for standard assumptions about clause structure and phonology within the HPSG framework. We demonstrate that argument composition is not an adequate explanatory device for Spanish VP-internal subjects, obviating the need for some modification to the standard account. Following the Bonami et al. (1999) account of a related French phenomenon, we outline a linearization analysis of Spanish clauses. However, this paper also briefly explores the possibility of using an alternative constituent-based approach to account for this data, as a matter for further research.

*Thanks to my informant M. Aránzazu Martín-Lozano for providing the original Spanish data in this work, and to Bob Levine, Detmar Meurers, Carl Pollard and Wesley Davidson for their comments and suggestions at all stages of this research. All remaining errors are my own.

Contents

1 Introduction

The term *postverbal subject NP* is used to refer to cases where the subject of a clause appears somewhere to the right of the finite verb that selects it. In Spanish, th subject NP may usually either precede or follow the finite verb, as illustrated by the examples in (1).[1] Note that the subject NP is shown in bold face, and the finite verb that selects the subject is underlined. We follow this convention throughout the paper to allow for easy identification of the two key elements.

(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.'

An *obligatory postverbal subject NP* refers to a case where the subject NP must occur to the right of the finite verb that selects it. Evidence of obligatoriness is illustrated by a minimal pair of clauses that differ only in the position of their subject NPs: the postverbal subject is grammatical, the preverbal subject is not. An example of an obligatory postverbal subject is shown in (2).

(2) a. ¿Qué querían **esos dos?**
what wanted those two
'What did those two want?'

b. * ¿Qué **esos dos** querían?
what those two wanted

An adequate HPSG account of Spanish clause structure must include (i) a characterization of the syntactic conditions under which subject NPs are obligatorily postverbal, (ii) a description of possible word orders of the postverbal subject NP with respect to other elements in the clause, and (iii) mechanisms and constraints which serve to license the necessary restrictions on the position of Spanish subject NPs. This paper is concerned with these issues, comparing the Spanish data to a comparable French phenomenon, and drawing on the (Bonami et al., 1999) account of that phenomenon to develop an account of Spanish postverbal subject NP clauses.

Part I addresses the conditions under which postverbal subject NPs are obligatory. Section 2 is concerned primarily with a description of those conditions. Preliminary data, much of it described in Torrego (1984), is presented in section 2.1. Predictions made by the current HPSG treatment of unbounded

[1] Describing the full inventory of Spanish clause types and the conditions (including pragmatic factors) that determine subject placement is beyond the scope of this paper. We will not discuss the conditions under which a preverbal subject may be obligatory.

dependencies are discussed in section 2.2; the Spanish facts are shown to be consistent with those predictions. In section 2.3, we introduce new data pertaining to obligatory postverbal subjects, arguing that a straightforward account of obligatoriness is possible given certain assumptions about Spanish interrogative clauses.

In section 3, we discuss a comparable phenomenon: the French optional postverbal subject NP in an extraction environment, as it has been described and analyzed in Kayne and Pollock (1978), henceforth K&P, and Bonami et al. (1999), henceforth BGM. In section 3.1 we compare the Spanish and French phenomena, noting that the French facts are also consistent with the predictions outlined in section 2.2. Section 3.2 covers the HPSG account of optional postverbal subjects developed by BGM.

Finally, we present our treatment of Spanish obligatory postverbal subjects in section 4, comparing the constraint introduced here to enforce obligatory postverbal subjects with the constraint formulated by BGM to license optional postverbal subjects.

Part II covers the possible word order associated with postverbal subject NPs in Spanish. In section 5 those word orders are described, while the problem they pose for an HPSG account making traditional assumptions about clause structure and phonology (Pollard and Sag, 1994) is explained in section 6. In this section we outline several conceivable solutions. We explore the first of these possible solutions, argument composition, in section 7, showing that predictions made by this analysis are not borne out by the Spanish word order facts. In section 7.1 we provide some background information which helps to make clear the predictions presented in 7.2. Also in this section, we provide the data which obviates the inadequacy of argument composition as an explanatory device for some VP-internal subject NPs, and hence, for Spanish postverbal subjects in general.

The next three sections deal with linearization—a second possible solution to the problem of word order in inverted clauses. Section 8 is a general discussion of linearization; section 8.1 presents the approaches developed by Reape (1994) and Kathol (1995), while section 8.2 outlines the details of the treatment assumed in this account. Section 9 discusses word order facts in French inverted clauses, and the linearization account developed by BGM to cover that data. The data is summarized in section 9.1, while the treatment is summarized in section 9.2. Constraints necessary to enforce the appropriate word orders in Spanish inverted clauses are introduced in section 10.

Part I
Obligatoriness

This part of the paper is concerned solely with the conditions under which the occurrence of a postverbal subject is obligatory, ignoring entirely the question

of how word order is licensed. In other words, pretend for the time-being that the term *postverbal* is an arbitrarily chosen name; in fact, we will license obligatoriness by means of a feature value that must independently be tied to word order in part II.

The key point to be argued for here is that a straightforward expression of the factors conditioning obligatory postverbal subjects is possible if we assume that not all Spanish *wh*-interrogative clauses with clause-initial *wh*-phrases are unbounded dependency constructions. This assumption provides an explanation for differences in meaning between certain interrogative clauses which differ superficially only in whether the subject preceded or follows the finite verb selecting it. The account which follows from this assumption holds that obligatory postverbal subject NP is a phenomenon associated with those *wh*-interrogatives which are head-filler constructions, and does not depend, for example, solely on some inherent property of the *wh*-phrase itself.

Furthermore, predictions made by the HPSG treatment of head-filler constructions (Pollard and Sag, 1994) are shown to be consistent with facts about both the Spanish and French postverbal subject phenomena.

2 Spanish obligatory postverbal subjects

2.1 Data

According to Torrego (1984), postverbal subject NP is obligatory in embedded (3c-3d) and non-embedded (3a-3b) *wh*-interrogative clauses with preposed *wh*-phrases.

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

b. * ¿Qué **esos dos** querían?
what those two wanted

c. No sabía qué querían **esos dos**
not knew what wanted those two
'I didn't know what those two wanted'

d. * No sabía qué **esos dos** querían
not knew what those two wanted (Torrego, 1984)

A postverbal subject is obligatory regardless of the position of the gap in the clause. That is, the clause-initial filler may correspond to a missing subject, complement, or modifier, in either the main clause or an embedded clause.[2]

[2] An obvious exception the case in which the preposed *wh*-phrase is itself the subject of the interrogative clause, so that a postverbal subject is by definition impossible (to say nothing of obligatory).

(4) a. ¿Con quién vendrá **Juan** hoy?
with whom will-come Juan today
'With whom will Juan come today?'

b. ¿Con quién vendrá hoy **Juan?**
with whom will-come Juan today
'With whom will Juan come today?'

c. * ¿Con quién **Juan** vendrá hoy?
with whom Juan will-come today

d. ¿Quien pensó **Juan** [que $_i$ se comió la manzana]?
who thought Juan that CL(to-him) ate the apple
'Who did Juan think ate the apple?'

e. * ¿Quien **Juan** pensó [que $_i$ se comió la manzana]?
who Juan thought that CL(to-him) ate the apple

The examples in (3) involve complement gaps. Examples of adjunct gaps and embedded subject gaps are shown in (4a)-(4c) and (4d)-(4e), respectively.

As stated above, the a postverbal subject is obligatory only in the presence of a *wh*-phrase that is *preposed*. In a *wh*-interrogative clause where the *wh*-phrase occurs *in situ* (i.e. an echo question), a preverbal subject is grammatical (5a). In fact, in this case a postverbal subject is not allowed, as shown by (5b) and (5c).

(5) a. **Marta** quiere qué?
Marta wants what (Torrego, 1984)
'Marta wants what?'

b. Quiere **Marta** qué?
Marta wants what

c. Quiere qué **Marta?**
Marta wants what

Furthermore, relative clauses, which can also contain preposed *wh*-phrases, do not condition an obligatory postverbal subject NP (6).

(6) a. el libro que escribió **Juan**
the book that wrote Juan
'the book that Juan wrote'

(i) a. ¿**Quién** le dio el libro a Juan?
who to him gave the book to Juan
'Who gave the book to Juan?'

b. * ¿Le dio **quién** el libro a Juan?
to him gave who the book to Juan

b. el libro que **Juan** escribió
the book that Juan wrote
'the book that Juan wrote'

As shown by (6a) and (6b), both preverbal and postverbal subjects are grammatical in a relative clause. Thus, a clause-initial interrogative *wh*-phrase conditions the presence of an obligatory postverbal subject, but a clause-initial relative *wh*-phrase does not.

2.2 Predictions made by HPSG about extraction-sensitive phenomena

In HPSG, the distinction between interrogative and relative *wh*-phrases is encoded via the *nonlocal* features QUE and REL, respectively (Pollard and Sag, 1994). A simplified version of the lexical entry for an interrogative pronoun, characterized by a non-empty QUE value and an empty REL value, is shown in figure 1.

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC|CAT|HEAD} & \textit{noun} \\ \text{NONLOC} & \begin{bmatrix} \text{REL} & \langle\ \rangle \\ \text{QUE} & \textit{nelist} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 1: Distinctive features of an interrogative pronoun lexical entry

A relative pronoun, on the other hand, is characterized by a non-empty REL value and an empty QUE value, as shown in figure 2.

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC|CAT|HEAD} & \textit{noun} \\ \text{NONLOC} & \begin{bmatrix} \text{REL} & \textit{nelist} \\ \text{QUE} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2: Distinctive features of a relative pronoun lexical entry

Information about the filler in filler-gap constructions is encoded at every step along the filler-gap (or extraction) pathway via the list-valued SLASH feature. Since SLASH lists contain *local* objects, however, the structure-sharing of information between fillers and gaps is limited to their *local* values, and only the *local* specifications of the filler are available along the filler-gap pathway. This means that the REL and QUE values of the filler are not visible along the filler-gap pathway, since they are *nonlocal* features.

More information about the filler is available at the level it is realized, since it is a non-head daughter (a *sign*). Crucially to the analysis of obligatory postverbal subjects, a constraint on a head-filler phrase can be formulated relating HEAD feature specifications to the to the NONLOCAL feature specifications of

the filler daughter. Since the Head Feature Principle (HFP) enforces structure-sharing of HEAD feature values between a phrase and its head-daughter, the HEAD feature specifications of a maximal projection and all intermediate levels are guaranteed to be those of the lexical head, in this case, the finite verb of the matrix clause. Thus, the feature geometry of HPSG allows the positing of head-filler constructions in which the presence of a filler which is an interrogative *wh*-phrase, a relative *wh*-phrase, or neither, determine *head* properties of the finite verb.

On the other hand, this set-up does not permit the encoding of a relationship between the *nonlocal* properties of a filler and *head* properties of an embedded clause along the extraction pathway. As mentioned above, the only information about a filler which is available to a finite verb heading a gapped embedded clause is that which is encoded in the SLASH list. For example, if we propose to encode the relative order of a subject with respect to the finite verb (preceding or following) as a HEAD specification, we cannot impose a constraint on embedded clauses which relates subject position to the presence of a particular class of filler higher up.[3]

Thus, while researchers have noted a number of syntactic phenomena that mark each step along an extraction pathway by means of the morpho-syntactic properties of finite verbs (Hukari and Levine, 1995), the treatment of unbounded dependencies outlined above predicts that there are certain restrictions on the kinds of extraction-sensitive phenomena we should expect to find. In short: a syntactic phenomenon involving the morpho-syntactic properties of finite verbs cannot both (i) be associated with every step along an arbitrarily long extraction pathway and (ii) depend upon the presence of a filler with a non-empty QUE or REL value at the top of that pathway. For instance, we might find a language where every finite verb along an extraction pathway must have a postverbal subject, but this restriction would have to hold for *any* head-filler construction, not only *wh*-interrogatives, for example. The existence of data contradicting this generalization would presumably signal the need to re-evaluate some aspect of the feature geometry.

However, evidence from Spanish obligatory postverbal subjects (and French optional postverbal subjects, to be discussed later) only serves to confirm the present HPSG treatment of UDCs. In an embedded finite clause along an extraction pathway, the subject does not need to be postverbal, as illustrated by (7).

(7) a. ¿Qué$_i$ pensaba **tu padre** [que **Alberto** le daría _$_i$ a Juan]?
what$_i$ thought your father that Alberto to-him would-give _$_i$ to Juan

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

[3]By class of filler, we mean in particular: relative *wh*-phrase, interrogative *wh*-phrase, or non-*wh*-phrase.

b. * ¿Qué **tu** **padre** pensaba [que **Alberto** le daría a
what your father thought that Alberto to-him would-give to
Juan]?
Juan

In (7a), the subject of the matrix clause must be postverbal (as shown by (7b)), but the subject of the embedded clause, from which *qué* is extracted, need not be. However we formulate the constraint licensing obligatory postverbal subjects, it must apply only to the level at which the filler is realized: the matrix clause.

2.3 New data: refining the relationship between obligatoriness and *wh*-interrogatives

All the evidence discussed so far points to an account of obligatory postverbal subjects in which it is not necessary to make reference to extraction (i.e. non-empty SLASH values or head-filler constructions) at all. Might it not simply be the case that a clause whose non-head daughter has a non-empty QUE value must also be headed by a finite verb whose subject is postverbal? The fact that the only clauses generally assumed to have such a daughter are head-filler constructions is not relevant to the formulation of a minimally descriptive constraint.

However, the generalization that all clause-initial interrogative *wh*-phrases condition an obligatory postverbal subject is incorrect, since clause-initial *por qué* and *cómo* both co-occur with preverbal subjects. Grammatical examples of *por qué* interrogatives involving both preverbal and postverbal subjects are shown in (8).

(8) 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
'What was Juan's reason for saying that he had given the book to Alberto?'
'What did Juan say was his reason for giving 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
'What did Juan say was his reason for giving the book to Alberto?'

Note that uninverted *wh*-interrogatives have a more restricted interpretation than their inverted counterparts. The reading of (8a) 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 presence of a postverbal subject NP. (That is, this reading is unavailable in (8b), in which the subject of the matrix clause is not postverbal.)

A straightforward account of this fact is possible given the following assumptions:[4]

[4]It is not possible to explain the obligatory/optional postverbal subject distinction in *wh*-

1. *Por qué* can combine with an ungapped clause to form an interrogative clause (which is not an instance of a head-filler construction).

2. An obligatory postverbal subject NP is an extraction-sensitive phenomenon in the sense that only the subject of an interrogative which is a head-filler construction must be postverbal. Subject position in interrogatives which are not head-filler constructions is unconstrained.

Note that these assumptions are consistent with the data, in particular the restrictions on possible meanings of uninverted interrogatives, but not necessary conclusions in light of that data. We lack independent evidence which either supports or contradicts the claim that *por qué* and *como* can attach to ungapped clauses, or that interrogatives with preverbal subjects are not filler-head phrases. In the absence of contradictory evidence, however, this approach is attractive for its explanation of the otherwise mysterious difference in meaning between (8a) and (8b), and because it allows a straightforward characterization of the syntactic environment conditioning the obligatory postverbal subject.

3 French optional postverbal subject NP (in extraction environments)

3.1 Data

In French, postverbal subject NPs are never obligatory, nor are they generally optional as is the case in Spanish. Rather, a postverbal subject is optional in several restricted environments, only one of which will be reviewed here, since it is the optionality of postverbal subjects in an extraction environment which is directly relevant to the Spanish data. We refer to the French phenomenon commonly known as *stylistic inversion*. According to K&P, subject NPs in French may optionally appear to the right of the verb in those clauses

interrogatives in terms of the argument/non-argument status of the *wh*-phrase. Both *por que* and *como* are normally analyzed as non-arguments, but so are *cuando* and *donde* which do require the presence of a postverbal subject (i).

(i) a. ¿Donde le dio **Juan** el libro a Alberto?
where clitic gave Juan the book to Alberto
'Where did Juan give the book to Alberto?'

b. * ¿Donde **Juan** le dio el libro a Alberto?
where Juan clitic gave the book to Alberto

c. ¿Cuando le dio **Juan** el libro a Alberto?
when clitic gave Juan the book to Alberto
'When did Juan give the book to Alberto?'

d. * ¿Cuando **Juan** le dio el libro a Alberto?
when Juan clitic gave the book to Alberto

characterized by preposed *wh*-phrases, namely non-embedded (9a) or embedded (9b) *wh*-interrogatives, and relative clauses (9c). BGM point out that, in fact, a postverbal subject NP is optional in "all well-known extraction contexts in French: relatives, *wh*-interrogatives or exclamatives, clefts, PP topicalizations."[5]

(9) a. Quand partira **ton ami?**
when will-leave your friend
'When will your friend leave?'

b. Je me demande [quand partira **ton ami**].
I wonder when will-leave your friend
'I wonder when your friend will leave.'

c. L'homme [avec lequel est sortie **Marie**] s'appelle Jacques.
the-man with whom has left Marie is-named Jacques
'The man who Marie has left with is named Jacques.'

In each of the examples in (9), the subject NP appears to the right of the finite verb. In clauses that do not involve extraction, a subject NP appearing to the right of the verb is ungrammatical. Each of the ungrammatical examples in (10) involves a postverbal subject in a non-extraction environment: a non-*wh*-interrogative clause (10a), a *wh*-interrogative with an *in situ wh*-phrase (10b and 10c), and a complement clause embedded within a declarative (10d).

(10) a. * Partira **ton ami?**
will leave your friend
'Will your friend leave?'

b. * Partira **ton ami** quand?
will-go your friend when
'When will your friend leave?'

c. * Partira quand **ton ami?**
will-go when your friend
'When will your friend leave?'

d. * Marie pense que a crié **Pierre.**
Marie thinks that has yelled Pierre
'Marie thinks that Pierre has yelled.'

Thus, in French as in Spanish, postverbal subjects (optional in this case) are associated with head-filler constructions. In French, however, *nonlocal* information about the filler is absent from the conditioning environment. As discussed

[5] Unfortunately, BGM do not provide examples of postverbal subjects in each of the cited extraction contexts, restricting their examples to relative clauses.

in section 2.2, HPSG's feature geometry predicts that the the postverbal positioning of subjects is a reflex which could be realized along an entire extraction pathway, since a non-empty SLASH value is sufficient to permit a postverbal subject in any clause.

As it turns out, a postverbal subject is also grammatical in an embedded gapped clause when the extracted *wh*-phrase is realized as a filler higher up, as shown in (11).

(11) Avec qui croit-elle qu'a soupé **Marie?**
with whom thinks-she that has dined Marie
'Who does she think that Marie has dined with?'

Not surprisingly, if there is no extraction from an embedded clause, but a gap site higher up, then a postverbal subject in that embedded clause is ungrammatical (12).

(12) a. Qui a dit que **Paul** pleure?
who said that Paul is crying
'Who said that Paul is crying?'

b. * Qui a dit que pleure **Paul?**
who said that is crying Paul
'Who said that Paul is crying?'

It is not possible, however, to see a chain of postverbal subjects along an entire extraction pathway, since BGM report that postverbal subjects are ungrammatical in clauses which contain complement clauses.[6]

(13) a. * l' étudiant [à qui$_i$ disait $\__i$ **Marie** [qu' elle ne viendrait pas]]
the student to who said Marie that she CL come not
'the student to whom Marie said that she would not come'

[6] K&P are more equivocal about whether such examples are always ungrammatical, claiming that some examples involving a postverbal subject in a inverted matrix clause with a clausal complement are merely unnatural. This holds whether the gap is local or located in the embedded clause.

(i) a. ? A quoi voit **Luc** [que Jean est venu]?
at what does see Luc that Jean has come
'What makes Luc say that Jean came?'

b. ? Avec qui$_i$ a prétendu **Marie** [que sortirait $\__i$ **Jean]?**
with whom did claim Marie that would leave Jean
'Who did Marie claim that Jean would leave with'

b. * le livre [qu'$_i$ avait cru [que Jean écrirait $\__i$] **mon éditeur**
the book that had thought that Jean write my publisher
parisien
Parisian]

'the book that my Parisian publisher thought that Jean would write'

In other respects, the syntactic environment conditioning the optionality of postverbal subjects is more complex than the conditions restricting Spanish subject placement. K&P report some additional restrictions on the kinds of complements a verb may have in order for a postverbal subject NP to be grammatical. For instance, in a footnote, they suggest that postverbal subjects are not possible in clauses with non-extracted NP complements (14).[7]

(14) * Comment sait **Paul** tout cela?
how knows Paul all this

'How come Paul knows all this?'

Furthermore, BGM discuss the properties of postverbal subjects themselves, concluding that they are, in fact, accusative NPs. I will not summarize that evidence here.

3.2 The BGM account: licensing optionality

To account for the data summarized above, BGM first introduce to the signature two subtypes of HEAD feature value *verb*: *non_inv_vb*, referring to non-inverted verbs, and *extr_inv_vb*, referring to inverted verbs in extraction environments.[8] A word has a HEAD value of type *extr_inv_vb* only if it has a non-empty SLASH value, it selects an accusative subject, and any complement it selects has a HEAD value of type *nonfinite*. Although BGM do not provide or discuss their *head* sub-hierarchy in any detail, they state that this last specification constrains the word to select "no finite sentence complement." This constraint is shown in figure (3).[9]

[7] No indication is given whether a subject occurring to the right of the direct object would be acceptable or not.

[8] Given the disjunctive nature of the environments in which postverbal subjects are optional according to their analysis, BGM cannot use the more familiar boolean HEAD feature INV unless they want to introduce a disjunctive description in the consequent of this constraint. In other words, since postverbal subjects are an option in more than one syntactic environment, they can either (i) encode the property of having a postverbal subject with a single feature description, and license its occurrence disjunctively: "if a postverbal subject, then environment 1 or ... or environment n", or (ii) they can encode the property of having a postverbal subject disjunctively, with a separate feature description for each environment licensing the occurrence of an optional postverbal subject. BGM choose (ii), apparently allowing for the inclusion of additional *inv_vb* subtypes of *head*, to cover optional postverbal subjects in non-extraction environments. Specifically, BGM cite these additional contexts in which a postverbal subject NP is optional: "(ii) heavy NP inversion; (iii) inversion in spatio-temporally dependent clauses, instantiated in three contexts: time adverbials, subjunctive complements, and sentences with a thetic interpretation in a narrative." They restrict their account to inversion in an extraction environment.

[9] BGM use the *list*( ) notation to mean "a list all of whose elements are specified as..."

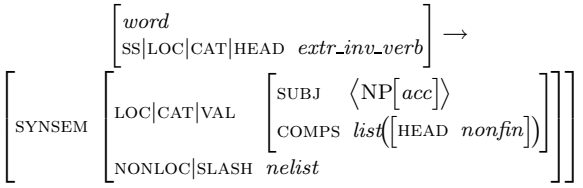

Figure 3: The BGM postverbal subject NP constraint for French

This constraint does not rule out a postverbal subject in a clause containing an *in situ* NP complement, something which K&P claim is ungrammatical, but is not mentioned by BGM. Furthermore, it is unclear whether or not the constraint glosses over restrictions regarding the acceptability of a postverbal in a clause with a sentential complement, given the difference in judgments between the two papers.

It should also be noted that the efficacy of this constraint depends on an account of extraction involving lexical amalgamation of SLASH values, as in the traceless account of Bouma et al. (2001), an approach fraught with semantic difficulties. In a configurational account of extraction like the trace-based account in Pollard and Sag (1994), the constraint on postverbal subjects must also be configurational: since the lexical head itself will not have a non-empty SLASH list, the constraint will have to apply to some point along the head path of the finite verb selecting the postverbal subject. It is far from obvious what the antecedent description in this constraint should be. A constraint like the one in figure 4, requiring a head-subject phrase with a HEAD value of type *extr_inv_vb* to have a non-empty SLASH list (and an accusative subject daughter) covers argument and VP modifier extraction.

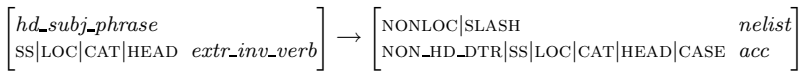

Figure 4: Configurational postverbal subject NP constraint for French

Extraction of a sentential modifier (if sentential modification exists and is relevant to the licensing of optional postverbal subjects) would not be covered by a constraint on head-subject phrases. In other words, such a configurational account on the lines sketched here will run into trouble if there exist extracted sentential modifiers which permit optional postverbal subjects. In the case that a constraint on head-subject phrases is adequate to restrict postverbal subjects to an extraction environment, an independent constraint, shown in figure 5 is necessary to encode restrictions on the selectional properties of a verb whose subject is postverbal, since the COMPS list is saturated at the head-subject phrase level.

In the next section, we posit a constraint licensing obligatory postverbal subject NPs in the appropriate syntactic environment, comparing the constraints

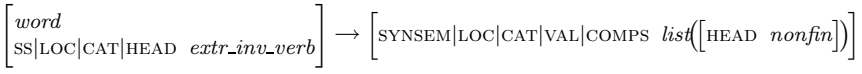

Figure 5: Constraint against sentential complements in inverted clauses

necessary to account for the French and Spanish phenomena of optional and obligatory postverbal subjects, respectively.

4 Licensing obligatoriness

As concluded in section 2.1, the syntactic environment conditioning an obligatory postverbal subject NP in Spanish is a head-filler construction whose filler daughter is an interrogative *wh*-phrase.

The structure of the BGM constraint is inadequate for the Spanish data because the conditional relationship between subject position and environment is different. That a postverbal subject is optional in French in a certain environment and ruled out elsewhere means that the environmental specifications are a necessary but not sufficient condition for the postverbal position. That is to say, every verb specified as *extr_inv_vb* will also have certain other specifications, but not every verb with those specifications need be specified as *extr_inv_vb*. In Spanish the situation is different. That a postverbal subject is optional in general but obligatory in a particular environment means that those environmental specifications are a sufficient but not necessary condition for the postverbal position. That is to say, all signs satisfying a certain description are required to contain a postverbal subject, but other signs may contain postverbal subjects as well.

We use the familiar boolean *head* feature INV(ERTED) to encode whether a verb selects a postverbal subject or not; an [INV +] specification on the verb signifies that its subject is postverbal. The constraint enforcing obligatoriness is given in figure 6.

Figure 6: Obligatory postverbal subject NP constraint

In short, the appropriate constraint for Spanish obligatory postverbal subjects holds that a head-filler construction with a filler daughter specified with a non-empty QUE value, must have the *head* specification [INV +]. Since this constraint applies to the phrase where a filler is realized, it is compatible with either a trace-based or a traceless account of extraction. It accounts for the data presented here provided one also assumes that *como* and *por qué* belong to a lexical class (sentential modifiers or complementizers, for example), whose

members combine with ungapped finite clauses, yielding phrases which are crucially not head-filler phrases.[10] It is still possible to refer to the general class of *wh*-interrogative clauses with the description given in figure 7.

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{HD\_DTR} & \left[\text{SYNSEM|LOC|CAT} \left[\begin{array}{ll} \text{HEAD} & \textit{verb} \\ \text{VAL} & \left[\begin{array}{ll} \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]\right] \\ \text{NON\_HD\_DTR|SYNSEM|NONLOC|QUE} & \textit{nelist} \end{array}\right]$$

Figure 7: Minimal description of a *wh*-interrogative clause

The distinguishing feature of a *wh*-interrogative clause, then, is that the non-head daughter has a nonempty QUE value.

In itself, the constraint in figure 6 does nothing to ensure that a subject NP shows up to the right of the verb that selects it; since the relationship between the feature specification $\left[\text{INF}\ +\right]$ and phonology has not yet been enforced. In the second part of this paper we turn to the matter of describing and enforcing possible word order in Spanish postverbal subject NP clauses.

Part II
Word order

In the first part of this paper, we addressed the conditions under which postverbal subject NPs are obligatory in Spanish. In this part, we discuss the possible word orders associated with postverbal subjects, whether obligatory or not. To the best of our knowledge, the restrictions on possible word orders are the same in either case. Specifically, we address the question of where exactly to the right of the finite verb a postverbal subject may occur.

The possible word orders associated with postverbal subjects pose a problem for an HPSG account making the following assumptions about a clause:

1. The finite verb realizes its subject, except in the case of subject extraction.[11]

2. Verb strings involve the selection of VP complements, and an embedded verb and its (non-subject) arguments form a constituent.

3. Phonological form is derived directly from constituent structure.

[10] Of course, these lexical items (or ones which sound like them) must also be able to occur as fillers in order to generate meanings which are dependent on extraction, as discussed in section 2.3. Note that this approach may lead to spurious ambiguity in the case that a particular instance of a *wh*-interrogative phrase can be analyzed either as extracted or occurring *in situ*.

[11] We use the term "X realizes Y" to mean that some projection of X has a non-head daughter Y.

Since the inverted subject of a finite verb may intervene between an embedded verb and one of its arguments (15), an adequate account of Spanish will need to dispense with one of the assumptions outlined above, and provide an alternative.

(15) ¿Qué quiere regalarle **Juan** a Pedro?
what wants to-buy-(to-him) Juan to Pedro
'What does Juan want to buy for Pedro?'

We might immediately question the validity of assumption (2), supposing instead that (15) involves a complex predicate, but, as will be shown in section 7, the structures licensed by argument composition are not adequate to resolve this problem.

Thus, an adequate HPSG grammar of Spanish will either need to appeal to a mechanism whereby phonological form is only indirectly derived from constituent structure (as in the linearization account we posit here), or allow for structures in which an embedded verb realizes the (postverbal) subject of the finite verb.

5 Data: Spanish inverted clauses

The subject NP selected by a finite verb may follow the verb's complements, as shown in (16).[12]

(16) ¿Qué quiere regalarle a Pedro **Juan?**
what wants to-buy-(to-him) to Pedro Juan
'What does Juan want to buy for Pedro?'

The subject NP may also occur between the finite verb and one of its complements, be it a VP (17a),[13] a PP (17b), an NP (17c), or a clause (17d).

[12]The ordering of a postverbal subject with respect to a sentential complement is more restricted. As shown by (i), a postverbal subject NP occurring after a clausal complement is questionable at best.

(i) ?? ¿Qué pensaba que Alberto vio **tu padre?**
what thought that Alberto saw your father
'What did your father think that Alberto saw?'

Since my informant's judgments are somewhat unclear, and to allow for the possibility that the restriction here may be prosodic and/or psycholinguistic, we assume in the analysis that (i) is grammatical.

[13]According to Torrego, both *ser* and *haber* disallow the occurrence of a postverbal subject between them and a VP complement, although the subject may be ordered within or following such a VP, as shown in (i). This suggests that *ser* and *haber* may form a complex predicate with the governed verb (Carl Pollard, p.c.).

(i) a. * ¿Qué le ha **tu padre** regalado a Alberto?
what to-him has your father given to Alberto

(17) a. ¿Qué quiere **Juan** regalarle a Pedro?
what wants Juan to-buy-(to-him) to Pedro
'What does Juan want to buy for Pedro?'

b. ¿Qué le prestó **Juan** a Alberto?
what to him lent Juan to Alberto
'What did Juan lend to Alberto?'

c. ¿A quién le prestó **Juan** el diccionario?
to whom to him lent Juan the dictionary
'To whom did Juan lend the dictionary?'

d. ¿Qué pensaba **tu padre** que Alberto vio?
what thought your father that Alberto saw
'What did your father think that Alberto saw?'

A subject NP may also occur within a complement of the finite verb that selects it when that complement is a VP, as seen by (18a). In fact, the subject NP may occur following a verbs string of any length. An NP subject may not occur within a clausal complement of the finite verb, however, as shown by (18b).

(18) a. ¿Qué quiere regalarle **Juan** a Pedro?
what wants to-buy-(to-him) Juan to Pedro
'What does Juan want to buy for Pedro?'

b. * ¿Qué pensaba que Alberto le daría **tu padre** a Juan?
what thought that Alberto to-him would-give your father to Juan
'What did your father think that Alberto would give to Juan?'

6 The problem for a traditional constituency-based account

Examples like (18a) and (19), in which the subject is realized within a complement VP, pose a problem for an HPSG account assuming that a finite verb realizes a subject, an embedded verb and its (non-subject) arguments form a

b. ¿Qué le ha regalado **tu padre** a Alberto?
what to-him has given your father to Alberto
'What has your father given to Alberto?'

c. ¿Qué le ha regalado a Alberto **tu padre?**
what to-him has given to Alberto your father
'What has your father given to Alberto?'

constituent, and 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.

(19) Quiere [comprar **Juan** el libro]$_{VP}$?
wants buy Juan the book
'Does Juan want to buy the book?'

Given a phonological form corresponding to the string in (20), where NP(Subj$_1$) is taken to be the subject of the finite verb V[fin]$_1$, and NP(DO$_2$) is taken to be a complement of the infinitive verb V[inf]$_2$, we see three alternative sets of assumptions about structure and phonology that might yield a descriptively adequate treatment of word order.

(20) V[fin]$_1$ V[inf]$_2$ NP(Subj$_1$) NP(DO$_2$)
quiere comprar Juan el libro

In the first case, it is possible to maintain the traditional constituent structure of a clause, as shown in figure 8, but that structure must be mapped to phonological form using the mechanisms of a linearization theory.[14] An analysis of word order in postverbal subject clauses along these lines is pursued in section 10.

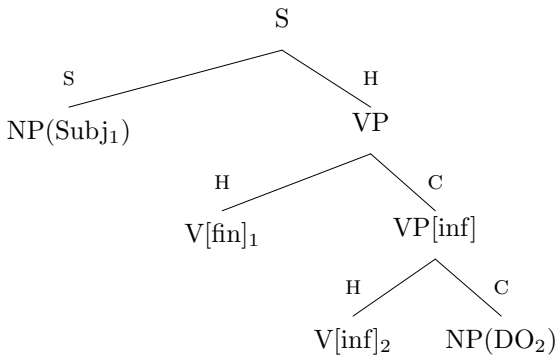

Figure 8: Possible structure 1 for V[fin]$_1$ V[inf]$_2$ NP(Subj$_1$) NP(DO$_2$)

In the second and third cases, it is possible to maintain a direct mapping of constituent structure to phonological form, but necessary to posit a different constituent structure, in which the VP[inf]$_2$, NP(Subj$_1$), and NP(DO$_2$) either (i) form a constituent which is a dependent of the VP[fin], as shown in figure 9, or (ii) do not form a constituent, and are individual dependents of the VP[fin], as shown in figure 10.

[14] It is not necessarily the case that Spanish has the oft-assumed VP[fin] as a constituent—we could also posit a flat structure here, the point being that constituent structure is not directly reflected in phonological form under this approach.

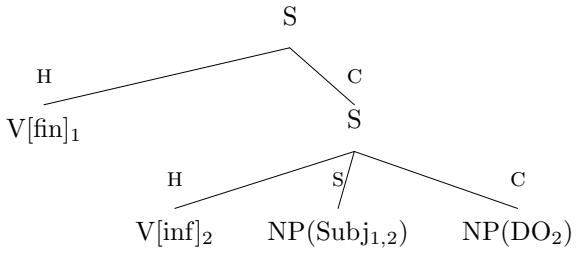

Figure 9: Possible structure 2a for V[fin]$_1$ V[inf]$_2$ NP(Subj$_1$) NP(DO$_2$)

The structure in figure 9 seems viable only if it can be shown that an NP subject occurring in an embedded VP is always the subject of both the finite verb and the embedded verb—in other words, the finite verb may select an inverted clausal complement. This works only if information about the subject is available to the finite verb, in order to account for agreement, case, etc.. For an example of data suggesting the need for such a structure in German, as well as mechanisms introduced within the HPSG framework to achieve the necessary sharing of information about the subject NP, see Meurers (1999). Preliminary data suggesting that this might be a fruitful approach in accounting for Spanish postverbal subjects will be discussed in section 11. While an in-depth exploration of this alternative to traditional assumptions about clause structure is beyond the scope of this paper, it is worth noting as a possibility for future research.

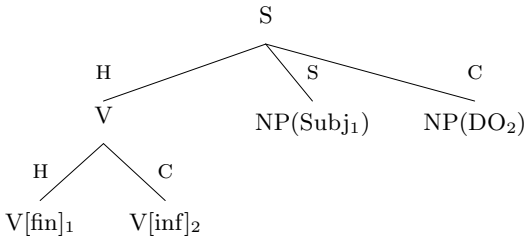

Figure 10: Possible structure 2b for V[fin]$_1$ V[inf]$_2$ NP(Subj$_1$) NP(DO$_2$)

The last alternative noted above, illustrated by the structure in figure 10, involves the formation of a complex predicate, and the inheritance of selectional properties, including complements, by the finite verb. In the next section, we explore an approach along these lines.

7 Argument composition and word order

To recapitulate, there are two basic word order facts to be accounted for here: (i) a postverbal subject NP may be interspersed with the complements of the matrix

verb, and (ii) the postverbal subject may be interspersed with the complements of a non-finite verb heading a VP selected as complement by the finite verb itself, or is embedded in a series of complement VPs. Accounting for (i) is relatively straightforward: presumably, the SUBJ list element of a finite verb marked as $\left[\text{INV } +\right]$ can be inserted into the COMPS list via a lexical rule.

The structure in figure 10, illustrating the generalization made in (ii), can be licensed by *argument composition* or *argument inheritance* (Hinrichs and Nakazawa, 1989). The term argument composition refers to cases in which verbs that normally select a VP complement select instead a lexical verb and its complements. Such verbs are often referred to as *restructuring verbs*. This variation in selectional properties is encoded by the inclusion of two distinct lexical entries, one derived from the other via a lexical rule like the one in figure 11.

$$\left[\text{SS|LOC|CAT}\ \left[\text{HEAD}\ \textit{verb}\ \text{VAL|COMPS}\ \left\langle\left[\text{LEX}\ -\ \text{CAT}\ \left[\text{HEAD}\ \textit{verb}\ \text{VAL}\ \left[\text{SUBJ}\ \textit{ne\_list}\ \text{COMPS}\ \langle\ \rangle\right]\right]\right]\right\rangle\right]\right] \mapsto$$

$$\left[\text{SS|LOC|CAT|VAL|COMPS}\ \left\langle\left[\text{LEX}\ +\ \text{COMPS}\ \boxed{2}\right]\right\rangle \oplus \boxed{2}\right]$$

Figure 11: Argument composition lexical rule

For example, in addition to the (inflected) base lexical entry for *quiero* shown in figure 12, which selects a VP complement, the argument composition lexical rule licenses the derived lexical entry shown in figure 13, which selects a lexical verb as well as its complements.

$$\left[\text{PHON}\ \textit{<quiero>}\ \text{SS|LOC|CAT}\ \left[\text{HEAD}\ \textit{verb}\ \text{VAL|COMPS}\ \left\langle\left[\text{LEX}\ -\ \text{CAT}\ \left[\text{HEAD}\ \textit{verb}\ \text{VAL}\ \left[\text{COMPS}\ \langle\ \rangle\right]\right]\right]\right\rangle\right]\right]$$

Figure 12: Selecting a VP argument

This lexical variation has the result that both both of the VP structures in figure (14) are possible. The structure on the left is headed by a verb which has not undergone argument composition; it selects a VP complement which itself consists of a lexical head and an NP complement. The structure on the right is headed by a verb which has undergone argument composition; it selects a lexical verb and its NP argument.

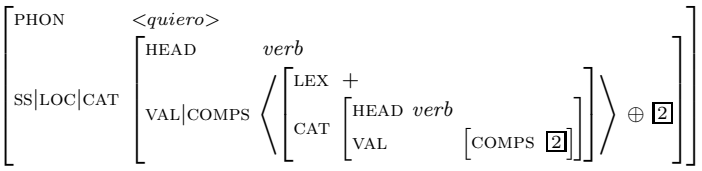

Figure 13: Selecting a verbal head + argument composition

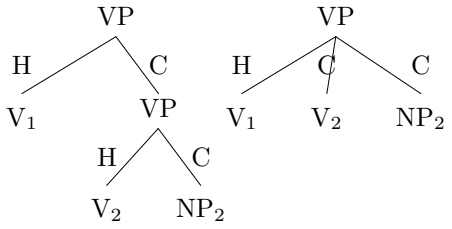

Figure 14: Two structures associated with restructuring verbs

7.1 Empirical evidence for argument composition

Syntactic tests suggested by Monachesi (1998), involving clitic attachment and a passive construction, are consistent with the existence of an argument composition lexical rule in Spanish. In her work on restructuring verbs and clitics in Italian, a language that shows a great deal of relevant similarity to Spanish, Monachesi (1998) discusses empirical evidence that certain verbs can head structures like both of those in 14, and uses argument composition to generate those structures.

Monachesi uses a variety of data to show that both such structures exist in Italian, all of which rely crucially on clitic placement; specifically, she associates the phenomenon of *clitic climbing* with argument composition. Much of this argumentation relies on movement; those tests are either unavailable for Spanish (pied-piping), suspect (RNR), or provide unclear results for Spanish (clefting). On the other hand, passive constructions provide robust results for both Italian and Spanish. Some background information about Spanish clitics will be presented, in order to later explain data pertaining to argument composition which involves cliticization.

7.1.1 Cliticization in Spanish

Pronominal clitics Spanish has a system of pronominal clitics marked for case, gender and number, and a small number of clitics serving other grammatical functions. Clitics occur in a predictable order, and there are restrictions on what combinations are possible. For an inventory of clitics and a summary of the system as a whole, see Zagona (2002) or Perlmutter (1971), for example. Here we have only a brief overview of cliticization in Spanish.

Complements (but not subjects) of a verb may be realized as a clitic in lieu of or in addition to an overt PP or NP complement. Indirect objects usually have a clitic copy associated with the verb, as in (21). Both the overt indirect object and its clitic copy are shown in boldface.

(21) Luis **le** envió la carta **a Pedro.**
Luis CL(to-him) sent the letter to Pedro
'Luis sent the letter to Pedro.'

Direct objects occur either as a full NP, or as a pronominal clitic (though not normally both simultaneously). Both possibilities are shown in (22).

(22) a. Juan vio **a Alberto.**
Juan saw to Alberto.
'Juan saw Alberto.'

b. Juan **lo** vio.
Juan CL-him saw
'Juan saw him.'

Pronominal clitics necessarily precede finite verbs, as shown in (23).[15]

(23) a. Juan **la** vio.
Juan CL(her) saw
'Juan saw her.'

b. * Juan vio**la.**
Juan saw-CL(her)

Pronominal clitics follow positive imperatives (24), infinitives (25), and present participials (26).

(24) a. Mira**la.**
look-CL(her)
'Look at her.'

b. * **La** mira.
CL-her look

(25) a. Voy a comprar**la.**
am-going to buy-CL(her)
'I am going to buy it.'

b. * Voy a **la** comprar.
am-going to CL(her) buy

[15] In Spanish orthography, preceding clitics are written as distinct lexical items, while clitics that follow the verb they are associated with are written as affixes.

(26) a. Estoy comprándo**la.**
am buy-CL(her)
'I am buying it.'

b. * Estoy **la** comprando.
am CL(her) buying

Clitics do not immediately precede or follow the past participial. Instead, clitics always precede all forms of *ser* and *haber*, which select the past participial. This pattern is shown in (27), with forms of *haber*.

(27) a. **La** he comprado.
CL(her) have bought
'I have bought it'

b. * He **la** comprado.
have CL(her) bought

c. * He comprado**la.**
have bought-CL(her)

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. 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*. Examples of this variability in clitic attachment are shown in (28).

(28) a. **Lo** estoy comprando.
CL(it) am buying
'I am buying it.'

b. Estoy comprándo**lo.**
am buying-CL(it)
'I am buying it.'

c. **Lo** quiero comprar.
CL(it) want to-buy
'I want to buy it.'

d. Quiero comprar**lo.**
want to-buy-CL(it)
'I want to buy it.'

As in Italian, however, clitics can not be "split" between verbs. That is, in those instances where clitics may occur with either a higher verb or an embedded verb, they must appear all on one or the other.

(29) a. Juan **te** **lo** quiere dar.
Juan CL(to-you) CL(it) wants to-give
'Juan wants to give it to you.'

b. Juan quiere dár**telo**
Juan wants to-give-CL(to-you)-CL(it)
'Juan wants to give it to you.'

c. * Juan **te** quiere dar**lo.**
Juan CL(to-you) wants to-give-CL(it)

Passive *se* In addition to the pronominal clitics discussed above, the clitic *se* can function as a passive marker on verbs with active morphology.[16] 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 (Martín-Lozano, Martín-Lozano). The following examples are provided by Martín-Lozano.

(30) 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'

Thus, in (30a) the subject is *el gobierno*, and the verb shows third person singular agreement; in (30b) the clitic *se* occurs with the verb and the verb shows third person plural agreement with *nuevas leyes*. Further evidence that the *nuevas leyes* should be analyzed as a subject is provided by the fact that it cannot be realized as a clitic (31); cliticization of subjects in Spanish is ungrammatical.

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

7.1.2 Constituent structure and cliticization

Restructuring verbs can occur with passive *se*, as shown in example (32).

(32) **Se** quieren quemar los libros.
CL-passive want(3pl) to-burn the books
'(Someone/they) want(s) to burn the books.'

[16] These passives are distinguished from the *periphrastic passive* construction by the absence of passive morphology, and the fact that they do not allow a *by*-phrase.

However, only the highest verb can occur with clitics in the presence of passive *se* (33a); attachment to the embedded verb is ungrammatical (33b), though fine in an active construction (33c).

(33) 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 dar**le** los libros.
CL want(3pl) give-CL(to-him) the books

c. Quiero dar**le** los libros.
want give-CL(to-him) the books
'I want to give him the books'

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 form of that verb. The grammaticality of (33a) supports the hypothesis that there is a lexical entry for *quieren* in which an NP is on its COMPS list. We assume here that this lexical entry is derived via argument composition, and that a passive lexical rule takes it as input in order to license (33a). Since clitic attachment to *quieren* co-occurs with *los libros* as a subject of *quieren*, the case in which we assume argument composition has applied, clitic climbing can be used as a diagnostic for argument composition.

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 any of its arguments as clitics (or alternatively, that the realization of any arguments as clitics by the verbal complement rules out argument composition). In cases where an embedded verb does realize clitics, as in (33c), 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.

In summary, the preceding data supports the association of clitic placement with the structures shown in figure (15).

Monachesi (1998) offers an explanation for the relationship between clitic attachment and possible structures. Very briefly, Monachesi claims that *complement cliticization* happens when a lexical rule removes a member of a verb's COMPS list and adds an element to the CLITICS set (a feature she introduces). That element is consistent with the complement's specifications in various ways, and is realized phonologically as an affix.

Thus, a clitic can only be attached to a verb if the corresponding complement phrase was on its COMPS list, and this can only be true for a restructuring verb in the case of argument composition. In Monachesi's account, argument composition stems from the application of another lexical rule, whereby a verb that selects for a VP may also select for a verb with that verb's own COMPS list

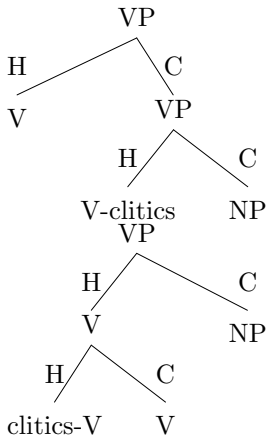

Figure 15: The interaction of cliticization and argument composition

appended, provided the complement verb's CLITICS set is empty. This last stipulation enforces the prohibition against split clitics, and is consistent with the evident relationship between clitic attachment and constituent structure illustrated in figure 15. Thus, her account formalizes the (apparently well-motivated) assumption that clitic climbing is a diagnostic for flat structure, and conversely, that clitic attachment to an embedded verb points to a VP constituent headed by that embedded verb, in which it realizes all of its remaining arguments, if any, as complements.

We turn now to predictions with respect to word order made by an argument composition account, depending on where clitics occur in a clause.

7.2 Predictions of argument composition about word order

A flat-structure account mapping constituent structure directly to phonological form can easily license the interspersal of a subject NP with the complements of the verb that selects it, but it cannot license the presence of the subject NP 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 (as signaled by clitic attachment to the higher verb), and ungrammatical where argument composition has not occurred (as signaled by clitic attachment to the embedded verb). In fact, as shown in example (34), both possibilities are fine.

(34) 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 [regalar**selo** **Juan** a Pedro]$_{VP}$?
want(3sg) buy-CL(to-him)-CL(it) Juan to Pedro
'Does Juan want to buy it for Pedro?'

Thus, even in the case that proponents of argument composition would assume involves a VP constituent (34b), the inverted subject can appear within the VP. Argument composition cannot explain this word order possibility. Any account insisting that example (34b) also exhibits a flat structure will have to account for why examples like (33b) are ungrammatical, losing the simplicity of a structural account like Monachesi's. Whether argument composition should be adopted to account for facts related to cliticization in Spanish is beyond the scope of this paper, but it is not readily apparent how it could be adequate to account for word order facts related to postverbal subject NP. We turn instead to an account involving linearization in the tradition of Reape and Kathol.

8 Linearization

8.1 Background: Reape and Kathol

As discussed by BGM, the idea behind linearization (as developed by Reape) is that the phonological form of a sign is not directly related to constituent structure, but instead related to the value of the feature DOM. DOM is a feature appropriate to objects of type *phrase*, and its value is a list of signs. "(T)he PHON value of a phrase is obtained by concatenating the PHON values of its DOM list, respecting order," as illustrated in figure 16 (Kathol, 1995). In addition, Reape introduces the feature UN(IONED), a binary feature whose value bears on how a daughter's DOM value contributes to the DOM value of the mother. Specifically, a daughter specified as [UN −] will be an element on the DOM list of the mother, while all signs on the DOM list of a daughter specified as [UN +] are also elements on the DOM list of the mother. The value of the mother's DOM list is gotten via the shuffle operation, as shown in figure 17.

$$\textit{phrase} \rightarrow \begin{bmatrix} \textsc{phon} & \boxed{1} \circ \ldots \circ \boxed{n} \\ \textsc{dom} & \left\langle \left[\textsc{phon}\ \boxed{1}\right], \ldots, \left[\textsc{phon}\ \boxed{n}\right] \right\rangle \end{bmatrix}$$

Figure 16: Derivation of a PHON value for a phrase from its DOMlist

$$\textit{phrase} \rightarrow \begin{bmatrix} \textsc{dtrs} & \boxed{0} \textit{list}\left(\left[\textsc{un}\ -\right]\right) \bigcirc \left\langle \begin{bmatrix} \textsc{un} & + \\ \textsc{dom} & \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} \textsc{un} & + \\ \textsc{dom} & \boxed{n} \end{bmatrix} \right\rangle \\ \textsc{dom} & \boxed{0} \bigcirc \boxed{1} \bigcirc \cdots \bigcirc \boxed{n} \end{bmatrix}$$

Figure 17: Constraint on the DOM value of a phrase, following Reape

One criticism of Reape's work among those that find the idea of linearization attractive, is that the value of the DOM feature should not be a list of signs, since signs carry more information than is needed. Pollard et al. (1993) suggests introducing the feature DOM_OBJ, and organizing the features PHON and SYNSEM underneath it. The value of DOM_OBJ is a type often called *dom_obj*, and the DOM list contains objects of this type. According to this approach, the basic architecture of a phrase might look like the (simplified) description shown in figure 18 (Kathol, 1995), and the basic architecture of a word might look like the description shown in figure 19 (Kathol, 2000). Notice that the DOM feature is appropriate for all signs (Pollard et al., 1993), not just phrases as it is for Reape.

$$\begin{bmatrix} \textit{phrase} & \\ \text{DOM\_OBJ} & \begin{bmatrix} \text{PHON} & \ldots \\ \text{SYNSEM} & \ldots \end{bmatrix} \\ \text{DOM} & \textit{list}\big(\textit{dom\_obj}\big) \\ \text{DTRS} & \ldots \end{bmatrix}$$

Figure 18: Proposed basic architecture of a phrase in post-Reape linearization

$$\begin{bmatrix} \textit{word} & \\ \text{DOM\_OBJ} & \boxed{1} \begin{bmatrix} \textit{dom\_obj} & \\ \text{SYNSEM} & \ldots \\ \text{PHON} & \ldots \end{bmatrix} \\ \text{DOM} & \langle\ \boxed{1}\ \rangle \end{bmatrix}$$

Figure 19: Proposed basic architecture of a word in post-Reape linearization

In his dissertation, however, Kathol himself does not adopt the feature DOM_OBJ, for reasons having to do with implications for the status of PHON values at intermediate phrasal stages. Briefly, nothing ensures that the PHON value (derived as in figure 16) of an embedded phrase is a subpart of the PHON value higher up, and in Kathol's estimation, "it seems preferable to try to ensure that the value for the PHONOLOGY feature is only computed if the corresponding constituent is domain-inserted or if the sign constitutes a complete, independent utterance." Towards the first case, he defines the relation *compaction*, which pairs a sign with the singleton list containing the appropriate *dom_obj*, as shown in figure 20.

For Kathol, then, while DOM is appropriate to all signs, and its value is a list of *dom_obj*s, there is no feature that has a *dom_obj* value. The basic structure of a word in Kathol (1995) (ignoring the feature TOPO, which is not relevant here) is illustrated in figure 21.

Moreover, Kathol (1995) does not adopt the feature UN.[17] Unlike BGM, he

[17]Where UN might be appropriate, Kathol utilizes the type hierarchy to make the necessary

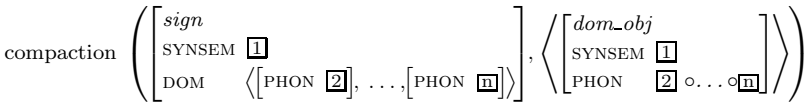

Figure 20: Kathol's constraint on the compaction relation

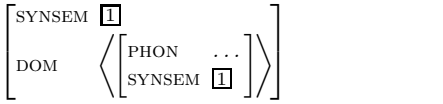

Figure 21: Basic architecture of a word in Kathol-style linearization

assumes binary-branching structures, and a DTRS feature that is not list-valued. Instead, sign-valued features such as HEAD_DTR and COMPL_DTR, which are presumably appropriate or not according to phrasal type, are organized under DTRS. He also abandons a list-valued SUBJ feature, reverting instead to a single SUBCAT list. Taken all together, these choices allow him to posit a single head-argument schema, which also specifies how the DOM list of the mother should relate to its daughters. This (slightly simplified) schema is shown in figure 22.

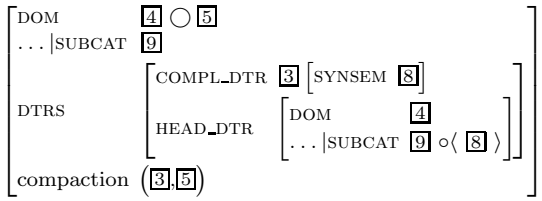

Figure 22: Kathol's binary head-argument schema

Thus, lexical heads and phrasal non-head daughters contribute domain objects to the DOM list of a phrase. In the case of non-head daughters, however, the appropriate domain object is gotten indirectly, via the compaction relation. Notice that there is no PHON value defined for *sign* as in traditional HPSG or Reape-style linearization. Although this does not seem to be discussed explicitly in Kathol (1995), we can presume that what an utterance sounds like is derived from the corresponding sign's DOM list via a relation like the one shown in figure 23, which associates a sign with the concatenation of the PHON values on its DOM list.

distinctions.

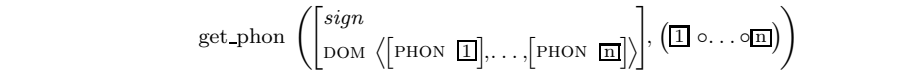

Figure 23: Relation deriving the phonological form of a sign via concatenation

8.2 Linearization in this account

For notational ease, we assume the DOM_OBJ feature rejected by Kathol (1995). As a result, the compaction relation is not needed. DOM is appropriate to all signs here, lexical entries are assumed to fit the description in figure 19, and the value of DOM_OBJ is a list of *dom_obj*s. We assume that the PHON value of a sign is computed as in figure 16 – the concatenation of all PHON values on that sign's DOM list, respecting order. In order to account for the ordering of external elements "within" other constituents, we also adopt the boolean UN feature, and assume it is a *head* feature.

We assume binary-branching structures (like Kathol), we are agnostic here about whether or not binary-branching has wider theoretical implications, but it is significant for the formulation of the constraints in figure 24. We maintain a list-valued SUBJ feature (*contra* Kathol, but more in line with current mainstream HPSG theory). Every phrase has the features HEAD_DTR and NON_HEAD_DTR.

Since we are using the UN feature, the value of which determines how the DOM list of the mother is related to its daughters, and having multiple VAL features makes it impossible to have a single head-argument schema, DOM specifications are not located in a schema as they are in Kathol (1995). That is, the inclusion of a domain object is treated the same in all cases in Kathol's approach, but depends on the value of UN in our approach. Thus, this approach requires two statements constraining the value of DOM, in contrast to Kathol's one. Furthermore, these constraints cannot apply to a single head-argument schema, as is the case in Kathol's approach, since the assumption of SUBJ and COMPS as valence features (instead of a single SUBCAT feature), necessitates two separate schemas licensing the realization of arguments.

Instead, we include two constraints on phrases, depending on whether the non-head daughter is specified as $[\textsc{un}\ +]$ or $[\textsc{un}\ -]$. These constraints are given in figure 24.

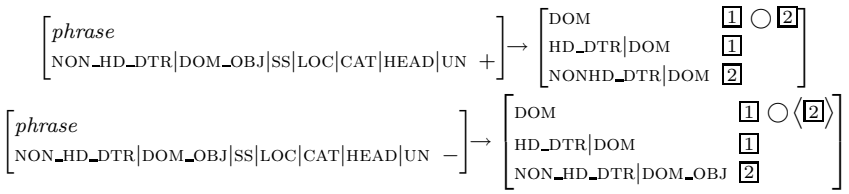

Figure 24: Constraints on the DOM value of a phrase

In the first case, the DOM list of the head daughter is shuffled with the DOM list of the non-head daughter, which allows ordering of other elements among the constituents comprising the non-head daughter. In the second case, the DOM list of the head daughter is shuffled with the singleton list containing the DOM_OBJ value of the non-head daughter. This is also referred to as *domain insertion*, and means that the non-head daughter is opaque to other elements in terms of word order. In both cases, the head daughter contributes its DOM list.

9 A linearization account of French inverted clauses

BGM offer a Reape-style linearization account of the word orders associated with French optional postverbal subject NP in an extraction environment, the phenomenon discussed in the first part of this paper. In the BGM account, the difference between an inverted and uninverted clause is located not in constituent structure, but in phonological form. In this section, we discuss the data presented by BGM pertaining to word order, and then summarize the mechanisms they use to account for that data.

9.1 Data: French word order

As previously noted, BGM distinguish between three varieties of postverbal subject NP in French, but restrict the scope of their investigation to postverbal subjects in extraction contexts. That investigation has two primary objectives – first, to describe the contexts in which a postverbal subject may occur, and second, to pay attention to where exactly the inverted subject shows up. Thus they present new data relevant to word order which are not captured by K&P, and it is that data we will summarize below. In short, they find that, as in Spanish, the inverted subject NP may occur after (or at the end of) the VP headed by the finite verb which selects it (as shown in (35a)), or between that finite verb and a complement (within the VP) as shown in (35b).[18]

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

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

BGM notice two further word order facts of import. The first is that an inverted subject may occur within an infinitival VP complement from which a

[18] Although BGM never point this out explicitly, the reader should note that in all data involving a verb and an *in situ* complement, that complement is a PP. In other words, they do not deal directly in their data or in their account of the data with the prohibition, mentioned by K&P, against postverbal subject when there is an *in situ* NP complement.

constituent has been extracted, no matter how deeply embedded it is, as shown in (36a) and (36b). Again, the ordering possibilities are similar to Spanish in this regard.

(36) a. le livre que pouvait recommander **le patron du labo** à cet
the book that could recommend the head of the lab to this
étudiant
student

'the book that the head of the lab could recommend to this student'

b. le livre que croyait pouvoir recommander **le patron du**
the book that thought can recommend the head of the
labo à cet étudiant
lab to this student

'the book that the head of the lab though he could recommend to this student'

Second, an inverted subject may never precede any verb; that is, although a subject NP may precede an NP or PP complement, it may not precede a complement VP as a whole. The relevant example, contrasting with (36b) above, is given in (37) below. French and Spanish differ here; subjects may intervene between verbs in Spanish, except where the first verb is a form of either *ser* or *haber*.

(37) * le livre que croyait **le patron du labo** pouvoir recommander
the book that thought the head of the lab can recommend
à cet étudiant
to this student

'the book that the head of the lab thought he could recommend to this student'

The only other constraint on possible orderings mentioned is a prohibition against an inverted subject NP occurring within a VP which is the complement of an object control verb, as illustrated by (38b). The occurrence of an inverted NP within a VP which is a complement of a subject raising or subject control verb is grammatical, as in (38a).

(38) a. le rôle bénéfique [que lui semblait [jouer **Pierre** dans ce
the role favorable that to-him seemed play Pierre in this
travail]$_{VP}$]
work

'the favorable role that Pierre to him seemed to play in this work'

b. * le livre [que m' a convaincu d' [offrir **mon libraire** à
the book that me has convinced DE offer my bookseller to
Marie]$_{VP}$]
Marie

'the book that my bookseller convinced me to offer to Marie'

That is, in the clause headed by the subject raising verb *semblait*, the inverted subject NP may occur within the complement VP headed by *jouer*. Note that in this case the subject *Pierre* is an argument of *jouer*. In the clause containing the object control verb *convaincu*, however, the inverted subject NP may not occur within the complement VP headed by *offrir*. In this case, where the subject NP *mon libraire* which shows up within the embedded VP is not an argument of its lexical head *offrir*, the sentence is ungrammatical.

This data suggests the following generalization about French postverbal subjects: a postverbal subject can only occur within a VP of which it is the subject. This generalization includes cases involving the raised or controlled subject of an embedded verb, as illustrated by the embedded VP-internal postverbal subject in (38a). The generalization crucially excludes cases involving embedded VPs whose (raised or controlled) subjects are realized as the object of a higher verb, in which case the postverbal subject is not the subject of the embedded VP, as illustrated in (38b).

The BGM account of this distinction reflected in (38) is stipulative, making no reference to the generalization above.[19]

9.2 The BGM linearization account

BGM formulate a default constraint requiring a phrase to be specified as $\left[\textsc{un}\ -\right]$, unless another constraint specifies otherwise. This constraint is shown in (25). A head-complement phrase with a HEAD value of type *extr_inv_vb* must be specified as $\left[\textsc{un}\ +\right]$, however. The BGM formulation of this constraint is shown in (26).[20]

$$\textit{phrase} \ \rightarrow \left[\textsc{ss|loc|cat|head|un}\ /-\right]$$

Figure 25: Default constraint for the feature UN

$$\begin{bmatrix} \textit{head\_comps\_ph} \\ \textsc{ss|loc|cat|head}\ \ \textit{extr\_inv\_vb} \end{bmatrix} \rightarrow \left[\textsc{ss|loc|cat|head|un}\ +\right]$$

Figure 26: Domain union constraint for VPs specified as *extr_inv_vb*

This is intended to account for data where an inverted NP shows up within a VP; a VP specified $\left[\textsc{un}\ +\right]$ allows the ordering of an NP within it, since LP rules apply to the individual elements on the DOM list of the VP, not the domain object of the VP. And, recalling that the complement VP of an object control verb must be [HEAD *non_inv_vb*], the default specification ensures that an inverted subject NP can not be ordered within such a VP.

[19] Preliminary data suggesting that the same generalization may be true of Spanish is discussed in section 12.

[20] BGM also include a third constraint pertaining to head-adjunct phrases, but since we restrict our attention to head-complement phrases in the relevant Spanish data, we do not include that constraint here.

From here, the burden of enforcing the proper word order possibilities rests with the LP statements. BGM give just two LP constraints on DOM values, shown below.

$$\left[\begin{array}{l} \textit{headed\_phrase} \\ \text{HD\_DTR } \boxed{1}\left[\text{SS|LEX } +\right] \end{array}\right] \rightarrow \left[\text{DOM } \boxed{1} \prec \left[\,\right]\right]$$

Figure 27: The BGM LP constraint requiring lexical heads to be domain initial

$$\textit{phrase} \rightarrow \left[\text{DOM } \left[\text{SS|LOC|CAT|HEAD } \textit{extr\_inv\_vb}\right] \prec \text{NP}\left[\textit{acc}\right]\right]$$

Figure 28: The BGM postverbal subject NP LP constraint

The LP statement shown in (27) specifies that a lexical head will always precede all non-heads. In general this ensures that when a $\left[\text{UN } +\right]$ DOM list is unioned into the DOM list of its mother, the lexical head of that mother precedes all else. Specifically, this ensures that in a structure with multiply-nested $\left[\text{UN } +\right]$ VPs, the verbs end up in the right order. The LP statement shown in (28) specifies that an NP[*acc*] is preceded by anything with HEAD value *extr_inv_vb*. This ensures that the subject occurs after all relevant verbs.

Given multiply nested slashed VPs which all have the same (inherited) subject, BGM need only ensure that the lowest verb has an NP[*acc*] on its SUBJ list, and that the highest verb is specified as *extr_inv_vb*, in order for the NP to be inverted. Exactly where it can show up is a function of which of the verbs along that pathway are *extr_inv_vb* (and thus [UN +]), instead of *non_inv_vb* (and thus [UN -]). At the topmost *non_inv_vb* VP, linearization boundaries are effectively set, so the inverted NP must occur before or after those boundaries. There is nothing explicitly prohibiting a *non_inv_vb* verb from taking a *extr_inv_vb* VP complement, but since the complement VP does not locally realize its subject, and its [UN +] specification will have no obvious effect in terms of linearization, one would never know. The problem is that, as we have just said, nothing requires that all nested VP complements be of the type *extr_inv_vb*. As Bob Levine initially pointed out to me (personal communication), this means that contrary to their claim, BGM do not rule out examples like the following:

(39) * le livre que croyait **le patron du labo** pouvoir recommander
the book that thought the head of-the lab can recommend
à cet étudiant
to this student
'the book that the head of the lab thought he could recommend to this student'

That is, imagine that *pouvoir* is in fact a *non_inv_vb*. None of the constraints presented by BGM prevent this. Then the LP statements given do not apply to the ordering of the subject NP with respect to the VP headed by *pouvoir*.

They could fix this problem by requiring that *extr_inv_vb* subcategorize only for *extr_inv_vb* VP complements, but then they lose the account of why inverted subjects of object control verbs can't occur between a selected verb and its complements, since that part of the account relies on the selected VP necessarily being *non_inv_vb*.

Despite the problems discussed here, the linearization account of word order in Spanish inverted clauses presented in the next section owes much to the BGM account of word order in French.

10 A linearization account of word order in Spanish inverted clauses

Once again, there are three basic word order facts to be enforced in a Spanish inverted clause. First, the subject NP must occur to the right of the finite verb that selects it (40). Second, the subject NP may be interspersed anywhere among the complements of the finite verb (40a-40b). Third, the subject NP may be interspersed with the complements of a verb heading a VP which is a complement of the finite verb itself, or is embedded in a series of complement VPs (40c). No other word orders are possible.

(40) 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.'

As a preliminary step, we assume the LP constraint formulated by BGM to ensure that lexical heads are domain initial. This constraint is shown again in figure 29.

$$\begin{bmatrix} \textit{headed-phrase} \\ \text{HD-DTR}|\text{DOM-OBJ}|\text{SYNSEM}|\text{LOC}|\text{LEX} \ + \end{bmatrix} \rightarrow \begin{bmatrix} \text{HD-DTR} \ \boxed{1} \\ \text{DOM} \quad \boxed{1} \prec [\,] \end{bmatrix}$$

Figure 29: The BGM LP constraint requiring lexical heads to be domain initial

Note that this constraint is relevant only in accounting for the order of complements with respect to the head that selects them.

The occurrence of the subject NP to the right of the finite verb that selects it is enforced by the LP constraint in figure 30, which encodes a general requirement that anything specified as $[\text{INV} \ +]$ precedes any NP on a DOM list. Crucially, this constraint does not apply to head-filler phrases. Thus, in a *wh*-interrogative, the *wh*-phrase need not follow the verb.

$$\textit{hd-val-phrase} \rightarrow \left[\textsc{dom} \left\langle \left[\textsc{synsem|loc|cat|head|inv} +\right] \prec \textsc{np} \right\rangle\right]$$

Figure 30: Postverbal subject NP LP constraint

Remembering that the head daughter of a phrase contributes its DOM list (and not its DOM_OBJ value) to the DOM list of the mother, this constraint licenses both possible word orders illustrated in (40b) and (40a).

As discussed in section 8, the feature UN is used to mediate the relationship between constituent structure and word order. In order to maintain the opacity of most constituents (that is, disallowing an outside element from being ordered within them), while allowing for the (optional) transparency of some, UN should be constrained to have a positive value only in a phrase matching some (possibly disjunctive) description. At the heart of this analysis is the assumption that the only non-head daughters which can be transparent to outside elements in terms of word order are VPs specified as $\left[\textsc{inv} +\right]$. The constraint necessary to enforce this restriction is provided in figure 31.

$$\left[\textsc{dom-obj|synsem|loc|cat|un} +\right] \rightarrow \begin{bmatrix} \textit{head-comps-ph} \\ \textsc{dom-obj|synsem|loc|cat|head|inv} + \end{bmatrix}$$

Figure 31: The relationship between INV and the UN feature

Thus, a complement VP which is specified as $\left[\textsc{inv} +\right]$ can, when also specified as $\left[\textsc{un} +\right]$, contribute the individual elements of its own DOM list, consisting of the lexical head and its complements, to the DOM list of a head-complement phrase in which it is a non-head daughter. The INV value of a complement VP is unconstrained—it is not related to the INV value of the verb that selects it, and the single constraint on INV applies only to verbs with a (particular kind of) clausal maximal projection.

To license the word order in (40b) it is necessary to assume that the infinitive complement VP is specified as $\left[\textsc{inv} -\right]$. Otherwise, the LP constraint in figure 30 requires the subject to occur somewhere to the right of the infinitive verbal complement.[21]

Finally, the constraints already introduced also license the problematic word order in (40c). Assuming that both the finite and infinitive verb are specified as $\left[\textsc{inv} +\right]$ and that the infinitive verb is specified as $\left[\textsc{un} +\right]$, the contribution of the finite VP to the DOM list of the head-subject phrase is a DOM list containing both verbal lexical heads and the NP complement of the infinitive verb. Domain-insertion of the subject's *dom_obj* into this list, respecting the LP constraint in figure 30, yields the problematic VP-internal word order.[22]

[21] Note the spurious ambiguity associated with (40a), depending on the INV specification of the infinitive VP.

[22] Of course, the same assumptions also license the subject-final word order—yet more spurious ambiguity.

11 "Unraised" subjects?

Linearization offers the means for an adequate description of the Spanish data investigated thus far, but we have not yet discussed the third structural possibility, shown again in figure 32, posited in section 6. In light of the data so far, a treatment of embedded VP-internal postverbal subjects along these lines, which we will refer to as an "unraising" account, has nothing to recommend it above a linearization approach, except for an oft-noted prejudice against linearization. Whether or not unraising structures offer a viable alternative to linearization in accounting for Spanish (or French) postverbal subjects is beyond the scope of this paper. However, we bring attention to it as an avenue for further research, and present data below, which, while preliminary, suggests that unraising may in at least one respect be superior to linearization in developing an adequate account of postverbal subjects.

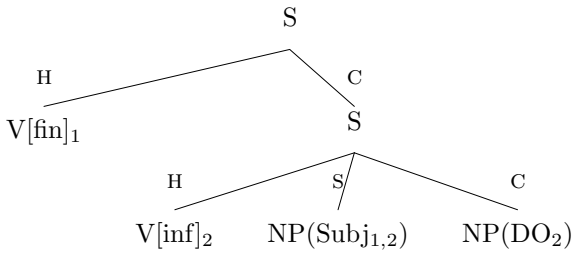

Figure 32: "Unraising" structure for V[fin]$_1$ V[inf]$_2$ NP(Subj$_1$) NP(DO$_2$)

As previously mentioned, one weakness of the BGM account of French postverbal subjects is that it offers no explanation for the fact that object control verbs do not allow the ordering of their postverbal subjects within their VP[inf] complement (in contrast to subject raising or control verbs, which do). If the same fact is true of Spanish, then this criticism can be leveled against the linearization account of Spanish postverbal subjects which has been posited here.

In fact, the preliminary data in (41) suggests that Spanish object control verbs may not, in fact, allow embedded VP-internal subjects. For example, the verb *vio* subcategorizes for a VP[inf] and its subject as complements (41a). That *Alberto* is a complement of *vio*, and not the subject of a clause headed by *comer*, is evidenced by the fact that *vio* can realize this argument as a clitic instead of an overt NP, as illustrated in (41b). Given that subjects cannot be realized as clitics, and that verbs can (apparently) only realize what is on their own COMPS list as clitics, we conclude that *vio* is an object control verb.

Crucially, and in contrast to data seen elsewhere in this paper, *Juan*, the subject NP of *vio*, cannot occur within the complement VP headed by *comer*.

(41) a. **Juan** vio a Alberto [comer la manzana]$_{VP}$.
Juan saw to Alberto eat the apple
'Juan saw Alberto eat the apple.'

b. **Juan** lo vio [comer la manzana]$_{VP}$.
Juan CL-him saw eat the apple
'Juan saw him eat the apple.'

c. * Vio a Alberto [comer **Juan** la manzana]$_{VP}$].
saw to Alberto eat Juan the apple

The point here is that *Juan* is not an argument of *comer*, and so the (ungrammatical) string *comer a Juan la manzana* cannot be analyzed as a clause. This data, then, lends itself to a hypothesis about what distinguishes grammatical VP-internal subjects from ungrammatical cases: a grammatical string consisting of a verb and a subject, followed by the verb's complements is, in fact, a clausal constituent. If the string cannot be analyzed as a clause (the subject NP is the subject of a higher verb, but not the verb immediately preceding it), then any utterance containing the string is ungrammatical. In other words, if the postverbal subject is not the (raised or controlled) subject of an embedded verb, then that postverbal subject cannot occur within the VP headed by that embedded verb.

If this hypothesis is borne out by further investigation, the approach developed in Meurers (1999) for German may prove fruitful here as well. In this approach, which maintains a direct mapping from constituent structure to phonology, verbs which ordinarily select a VP complement and inherit the subject specifications of that VP, may under certain conditions select a clause, while still licensing agreement with, and case assignment to, the subject, via the *head* feature SUBJ. In contrast to the list-valued valence feature SUBJ, the value of this feature (a *synsem*) makes information about the subject available at the level of a maximal projection, via the head path. The subject raising relationship between a lexical head and its verbal complement is licensed by structure-sharing of the *head* feature SUBJ, as is the converse "unraising" relationship. In an "unraising" structure, then, the finite verb of the matrix clause does not realize a subject: instead, its subject is realized by an embedded verb.

The "unraising" approach can account for embedded VP-internal postverbal subjects, but only when the individual elements of the VP, together with the postverbal subject, can be posited as a clause. This hypothesis about embedded VP-internal postverbal subjects is consistent with evidence from French, in which postverbal subjects cannot be realized within the complement VPs of object control verbs. (Again, the subject of the embedded VP in an object control construction is realized as a complement of the higher verb.) An example illustrating restriction is given again in (42).

(42) * le livre [que m' a convaincu d' [offrir **mon libraire** à
the book that me has convinced DE offer my bookseller to

Marie]$_{VP}$]
Marie

'the book that my bookseller convinced me to offer to Marie'

Evidence like this suggests that an account positing embedded VP-internal postverbal subjects as "unraising" constructions may have an explanatory advantage over the linearization accounts of postverbal subject NP constructions developed in this paper and by BGM. If further investigation is consistent with the preliminary data presented in this section, we would like to pursue an account along the lines of Meurers (1999) in future research.

12 Summary and Outlook

In this paper we have investigated on two aspects of the Spanish postverbal subject NP phenomenon, with a dual focus on (*i*) characterizing the syntactic environment that conditions an obligatory (rather than optional) postverbal subject, and (*ii*) describing the possible word orders associated with postverbal subjects.

Evidence pertaining to Spanish obligatory postverbal subject NPs and French optional postverbal subject NPs is consistent with the inclusion of REL and QUE as *nonlocal* features in the HPSG feature geometry. That is, relative *wh*-phrases are syntactically distinct from interrogative *wh*-phrases, and in a head-filler construction with one or the other as a filler, that distinction does not need to be visible along the extraction pathway in order to account for either of these phenomena. Furthermore, a straightforward account of obligatory postverbal subject NP is possible if it is assumed that not all *wh*-interrogative clauses are unbounded dependency construction.

The word orders associated with Spanish postverbal subjects cannot be explained by a traditional constituency-based HPSG approach, nor can they be explained by appeal to the flat structures licensed by argument composition. This leaves two remaining possibilities in accounting for word order facts. The first involves delinking constituent structure from phonological form, mapping one to the other via the mechanisms provided by linearization. Such an account has been presented here. The second possibility, which involves the realization of subjects by embedded complement VPs (an approach we have referred to as "unraising"), we have suggested as an avenue for future research.

References

Bonami, O., D. Godard, and J.-M. Marandin (1999). Constituency and word order in French subject inversion. In G. Bouma, E. W. Hinrichs, G.-J. M. Kruijff, and R. T. Oehrle (Eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, pp. 21–40. Stanford: CSLI Publications.

Bouma, G., R. Malouf, and I. A. Sag (2001). Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory 19*, 1–65.

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.

Kathol, A. (1995). *Linearization-based German Syntax*. Ph. D. thesis, The Ohio State University.

Kathol, A. (2000). *Linear Syntax*. Oxford University Press.

Kayne, R. and J.-Y. Pollock (1978). Stylistic inversion, successive cyclicity, and Move NP in French. *Linguistic Inquiry 9*(4), 595–621.

Martín-Lozano, M. A. Spanish direct objects and arbitrary *se*. ms.

Meurers, W. D. (1999). Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL) 43*, 173–226.

Monachesi, P. (1998). Decomposing Italian clitics. In S. Balari and L. Dini (Eds.), *Romance in HPSG*, pp. 305–357. Stanford: CSLI publications.

Perlmutter, D. M. (1971). *Deep and Surface Structure Constraints in Syntax*. Holt, Rinehart, and Winston.

Pollard, C. and I. A. Sag (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.

Pollard, C. J., R. D. Levine, and R. T. Kaspar (1993). Studies in constituent ordering: Towards a theory of linearization in Head-Driven Phrase Structure Grammar. Grant proposal to the National Science Foundation, Ohio State University.

Reape, M. (1994). Domain union and word-order variation in German. In J. Nerbonne, K. Netter, and C. Pollard (Eds.), *German in HPSG*, pp. 151–197. CSLI Publications.

Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry 15*(1), 103–129.

Zagona, K. (2002). *The Syntax of Spanish*. The Press Syndicate of the University of Cambridge.