NON-MATCHING CONCEALED QUESTIONS IN CATALAN AND SPANISH AND THE PROJECTION PRINCIPLE*

PAUL HIRSCHBÜHLER and MARÍA-LUISA RIVERO

In this paper, we study the properties of complements of verbs of knowledge/perception and saying, including free relatives, in Catalan and Spanish. Some of these structures, the non-matching concealed questions, violate the subcategorization frame(s) of the verbs they complement at the levels of D- and S-Structure, but are mapped into well-formed LFs. A language-specific rule gives them the type of LF that satisfies the lexical requirements of the subcategorized position they occupy. We explore the consequences of this proposal for the parametrization of the Projection Principle (Chomsky 1981).

1. THE PROJECTION PRINCIPLE AND CATALAN AND SPANISH VERBS OF KNOWLEDGE, PERCEPTION, AND SAYING

In his Lectures on Government and Binding, Chomsky proposes the Projection Principle. Intuitively, this principle states that the requirements of lexical items must be satisfied at each of the three syntactic levels, namely Deep Structure, Surface Structure, and Logical Form. Formally the principle is stated as in (1) (Chomsky 1981:38).

(1) Projection Principle:
In structures of the form \([\delta..\alpha..\beta...]\) or \([\delta..\beta..\alpha...]\)

(i) if \(\beta\) is an immediate constituent of \(\delta\) at \(L_i\), and \(\delta = \tilde{\alpha}\), then \(\alpha \theta\) marks \(\beta\) in \(\delta\)

(ii) if \(\alpha\) selects \(\beta\) as a lexical property, then \(\alpha\) selects \(\beta\) in \(\delta\) at \(L_1\)

(iii) if \(\alpha\) selects \(\beta\) in \(\delta\) at \(L_1\), then \(\alpha\) selects \(\beta\) in \(\delta\) in \(L_j\).

The term selects in this context refers to direct \(\theta\)-marking for subcatego-

* This work has been partially subsidized by Research Grant 410-80-0660 from the Social Sciences and Humanities Research Council of Canada. We thank Margarita Suñer for comments on a previous version.
rized positions ((i) in (1)), and indirect $\theta$-marking for nonsubcategorized positions such as the subject (i.e. (ii) in (1)). Thus, a verb directly selects the phrases it subcategorizes for (e.g. the direct object if a transitive verb), and may indirectly select the subject. By (iii), the projection principle requires that selection be satisfied at the three syntactic levels. If the Projection Principle is to be satisfied, English examples such (2a) must be analysed as in (2b) (Chomsky 1981:96ff),

(2) a. John believed Bill to be sad  
    b. John [\_ believed [\_ Bill to be sad]]

because the VP containing believe assigns a $\theta$-role to the subject John, and the verb assigns one to its sentential complement, but not to the NP Bill. This situation is clear at the level of LF, and by the Projection Principle, it should also be the case at S- and D- structures. The traditional Raising-into-Object analysis is excluded because at S-structure the verb believe would select the NP Bill, while it would not select it at the other two levels of representation.

In Catalan and Spanish, verbs of knowledge and perception (comprender ‘to understand’, imaginar ‘imagine’, observar ‘observe’, saber ‘to know’, veure/ver ‘to see’, etc.), verbs of saying such as dir/decir ‘to say’, and the preposition/conjunction segons/según ‘depending on, according to’ exhibit properties which are problematic for the Projection Principle in relation to complements that we term non-matching concealed questions, including free relatives. To see why, let us consider the subcategorization properties of those items.

The verbs subcategorize for NP, (3), and $\bar{S}$-complements which can be declarative sentences, (4), or indirect questions, (5);

(3) a. Verás el libro  ‘You-will-see the book’
    b. Veuràs el llibre

(4) a. Verás que trabaja muy bien  ‘You-will-see that he-works very well’
    b. Veuràs que traballa molt bé

(5) a. Verás de qué hemos hablado  ‘You-will-see of what we-have spoken’
    b. Veuràs de què hem parlat

They do not subcategorize for PPs, and (6) is ungrammatical as a result:

(6) a. *Verás del libro  ‘You-will-see of-the book’
    b. *Veuràs del llibre
Segons/según ‘according to, concerning’ offers a similar situation.

However, the same classes of verbs, as well as segons/según can take as complements headed relatives with an initial PP, (7), and free relatives with similar characteristics, (8).\(^1\)

(7) a. Observe en les dificultats que ens vam posar (Viaplana 1980)
    ‘See into which difficulties we have gotten’

   b. Sé con los problemas que se enfrentan
    ‘I know which problems they face’

(8) a. Observe en el que ens vam posar
    ‘See into what ourselves we have put’

   b. Sé con lo que se enfrentan
    ‘I know what they face’

As partially indicated in the glosses, the sentences in (7) are indirect questions; those in (8) are too, and they share their interpretation with the unproblematic Observe en què ens vam posar and Sé con qué se enfrentan respectively. We label the complements in (7) and (8) non-matching concealed questions, and consider them a subtype of so-called concealed questions (Baker 1972, Grimshaw 1979, Heim 1980).

The PP relatives can also have exclamative interpretations with the appropriate matrix verbs, as in (9) and (10). Whether as exclamatives or as indirect questions, they display the same properties with respect to the issues discussed in this paper, when in embedded position. Según/segons does not take complements of the exclamative type.

(9) Tendrias que aver vist
    a. en les dificultats que ens vam posar!
       ‘You should have seen the difficulties we have gotten into’
    b. del que (= de lo que/de el que) hem parlat a la conferencia!
       ‘What we discussed at the meeting’

(10) Imagínate
    a. con los problemas que se enfrentan
    b. de lo que han hablado
    Imagine
    a. the problems they face!
    b. what they discovered!
To see the problems that the above structures pose for the Projection Principle in a preliminary and informal way, let us take a look at the analysis proposed for (7) by traditional grammarians. Translated into current generative terminology, the assumption is that the preposition en in (7a) and con in (7b) is moved away from the relative phrase in COMP to a pre-antecedent position in the structures which also underly grammatical (11a) and (12a).

(11) a. Observa les dificultats en què ens vam posar
Observe the difficulties into which ourselves we-have gotten
'See the difficulties we have gotten into'

   b. Observa [ les dificultats [ [ [ [en] què] [ ens vam posar ]]]

\[
\begin{array}{c}
\uparrow NP \\
\tilde{S} \text{ COMP PP P} \\
\text{S}
\end{array}
\]

(12) a. Sé los problemas con que se enfrentan
I-know the problems with which themselves they-face
'I know the problems they face'

   b. Sé [ los problemas [ [ [ [con] que] [ se enfrentan ]]]

\[
\begin{array}{c}
\uparrow NP \\
\tilde{S} \text{ COMP PP P} \\
\text{S}
\end{array}
\]

Under this approach, the subcategorization properties of observar and saber in (11a) and (12a) are met at the level of D-structure by the NP-complement, in a situation parallel to (3). However, subcategorization is violated at S-structure if the verb has a PP-complement at that level (similar to (6)). This is contrary to part (iii) of the Projection Principle in (1).

In the following section, we discuss the properties of (7), (9a), and (10a) in detail. For the problems connected with a formalization of the traditional analysis, see footnote 2. In section 3 we turn to the free relatives in (8), (9) and (10b). In section 4 we discuss the different parameters involved in the Projection Principle.

2. THE ANALYSIS OF NON-MATCHING CONCEALED QUESTIONS

As we have seen in the previous section, the sentences in (7), (9) and (10) have a matrix transitive verb followed by a PP-complement the verb does not select in the sense of the Projection Principle in (1), and are a type of concealed question (or exclamative). Current proposals for relative structures and questions suggest analysis (13a-c-d) for the surface structure of (7a). A fourth possibility is to treat en les dificultats as a TOPIC along the lines of (13b). Spanish sentence (7b) can receive parallel treatments. Traces are omitted.
V\textsuperscript{s} observa\textsuperscript{C O M}P\textsuperscript{s} en les dificultats que ens vam posar

\begin{itemize}
\item[(b)]
V\textsuperscript{s} observa\textsuperscript{TOP} S\textsuperscript{COMP} en les dificultats que\textsuperscript{S} ens vam posar
\end{itemize}

\begin{itemize}
\item[(c)]
V\textsuperscript{e} observa\textsuperscript{PP} S\textsuperscript{COMP} en les dificultats que\textsuperscript{S} ens vam posar
\end{itemize}
Each one of these analyses raises different problems. We proceed to consider them in turn, and show that (13d) is the one to be preferred. The S-structure in (13d) violates the subcategorization properties of the matrix verb; we propose that (13d) correlates with a well-formed LF, and would violate the Projection Principle at SS, if the theory of selection behind the formulation in (1) includes categorial information, an issue we discuss in section 4 when we consider the parameters of the principle.

Consider the analysis in (13a); the COMP is doubly-filled by a non wh-phrase and the complementizer que, but there is no problem with respect to subcategorization, and no violation of the Projection Principle. Given that verbs of knowledge and perception subcategorize for S-complements, as seen in (4) and (5), tree (13a) is unproblematic at each level of representation, and satisfies condition (iii) of the Projection Principle. However, a decisive argument against (13a) can be developed on the basis of (14) and (15a), already documented in the Spanish of the 17th century, as seen in the often cited (15b):

(14) Observa en les dificultats en què ens vam posar
Observe in the difficulties in which ourselves we-have-put
(= 7a)

(15) a. Sé con los problemas con que se enfrentan
I know with the problems with which themselves they-face
(= 7b)

b. Dárte a entender, Panza, en el error en que estas
To-make-you understand, Panza, in the error in which your-are
'To make you understand, Panza, which error you made'
*Quijote I, 15*

Because of the presence of the PP relatives in *COMP*, *en què* in (14), *con que* and *en que* in (15a-b)) the preceding PP cannot be in *COMP* as well. Among other reasons, there cannot be two successive PPs in *COMP* as the result of *Wh*-movement, given that there is only one site in the embedded clause for the movement to originate from. We take the examples in (14) and (15) as an indication that a PP outside of the boundaries of $\bar{S}$ must be postulated, making those structures available for the examples in (7), (9), and (10) as well. Another argument against placing the first PP in *COMP* can be built on the basis of PP complements whose head noun is *manera*. Consider (16a-b).

(16) a. De la manera en què has has entrat és imperdonable
   Of the manner in which you have come-in is inexcusable
   'How you came in is inexcusable'

b. Compta, doncs, de la manera {com }
   Tell, therefore, of the manner {how, in which}
   vivien els jesuites a Bolònía (Calveras 1929:217)
   lived the Jesuits in Bolonia
   'So tell how Jesuits used to live in Bolonia'

c. Quan veig les dones, de la manera {com }
   When I-see the women, of the manner {how, in which}
   van a combregar, tota em despaciento
   go to take-communion, all myself I-get-upset

   (Calveras 1929:217)
   'When I see how women go to take communion, I get all upset'

Spanish examples are parallel in the relevant respects. Complements with *manera* require the preposition *de* as in *Has entrat d’aquesta manera* ‘You-came in of this manner’; but *manera* as an antecedent requires the *P en* in *COMP* in relatives as in *la manera en què has entrat* ‘the manner in which you came in’. This is exactly the distribution mirrored in (16): *en* appears in *COMP* with the relative pronoun, *de* appears elsewhere. We conclude again that the example in (16b) requires the postulation of a PP outside of the boundaries of $\bar{S}$, arguing indirectly against analysis (13a). The difference between (14) and (7a) can be attributed to the deletion of the PP-relative in *COMP*, and the surfacing of the complementizer *que*, in an analysis along the lines of Hirschbühler and Rivero (1981a), or to the movement of an empty category and the insertion of the complementizer
que, in more recent analyses of relatives (see Chomsky 1981 for instance) that dispense with deletion rules.

In Spanish, there are other relative-like structures which are interpreted as questions (or exclamatives): the exocentric relatives in (17). Again, their antecedent-like phrases (lo guapa in (19a), and lo elegantemente in (17a)) do not appear in COMP, as we shall see shortly. This situation provides indirect evidence against (13a) for (7), even though (7) and (17) cannot receive identical treatments. Later on we consider some differences between (7) and (17).

(17) a. Juan vió lo guapa que María era
    John saw the pretty that Mary was
    ‘John saw how pretty Mary was’
b. Juan vió lo elegantemente que vestían
    John saw how elegantly that they-dressed
    ‘John saw how elegantly they dressed’

Recently, there have been several proposals concerning the structures in (17) (Knowles 1978, Rivero 1980a, Plann 1980, García-Bellido 1981). In relation to the issues we are discussing now, Plann proposed to analyse lo guapa que and lo elegantemente que as strings in COMP, while the other authors place que in COMP and the preceding phrase in the antecedent position of a restrictive relative (i.e. outside of COMP), with different mechanisms that ensure that the adjective in (17a) agrees with the subject of the embedded clause.

In view of the example of coordination in (18), which we take from García-Bellido (1981), the second proposal (i.e. (19)) is preferable. We concentrate exclusively on whether the AP should share the COMP with the complementizer que or not.

(18) Juan vió lo guapa que María era y que Alicia parecía
    John saw the pretty that Mary was and that Alice seemed
    ‘John saw how pretty Mary was and Alice seemed (to be)’
Under the antecedent analysis sketched in (19), (18) is accounted for in a straightforward manner: there is one antecedent modified by conjoined Šs, each one of them introduced by the complementizer *que*. Under a COMP analysis along the lines of (13a), *lo guapa* would have to be deleted or interpreted in the second COMP, but *lo guapa* in the first COMP would not have scope over the second COMP, and could not license the appropriate gap in that position. In brief, constructions such as (17) and (18) are also best analysed with the phrase beginning with *lo* and preceding *que* outside of COMP. These constructions share with those in (7) the fact a) they have an indirect question interpretation, b) the initial phrase is obligatorily followed by a relative-like structure, c) this initial phrase appears not to satisfy the subcategorization requirements of the matrix verb. Under these conditions, conclusions based on the characteristics of (17) and (18) give indirect support against analysis (13a) for (7). In conclusion, the complements in (7) are not simply Šs.

The remaining structures in (13) share the desired feature of having *en les dificultats* external to the embedded Š. Let us now turn to (13b) which analyses (7) as a topicalized or left-dislocated structure.

The structures in (7) differ from topicalizations or left-dislocations in several respects:

a) In (7), the complementizer *que* follows the PP that should count as the embedded topic. In left-dislocations and topicalizations the opposite order obtain, as in (20).
The complementizer que cannot precede the PPs in (7).
is not problematic for the Projection Principle because verbs of knowledge and perception allow embedded topics (i.e. they may subcategorize for \( \bar{S} \) and also \( \bar{S} \), if the tree for a bona fide embedded topic is along the lines of (13b)).

Let us now turn to analyse (13c). This tree goes against the \( \bar{X} \)-schema because the dominating node of the relative-like structure NP and its head (PP) do not have the same categorial specifications. In Rivero (1980a), the Spanish sentences in (17) contained a neutralized [+N] node in antecedent position. Since Ns, As, and Advs share the [+N] feature, \( \text{lo guapa} \) in (17a) and \( \text{lo elegantemente} \) in (17b) can function as derived heads of the relative-like construction in a way which is compatible with \( \bar{X} \)-theory. This approach cannot be extended to (7), because no feature is shared by Ns and Ps. Garcia-Bellido (1981) proposes category-switching rules in the base for constructions of type (17). These rules are similar to the deverbalizing rules postulated by Jackendoff (1977) for a limited set of English constructions, such as gerundive nominals, and free relatives. In (17), the complement of \( \text{vió} \) is an NP, but the head of that complement is an adjectival or adverbial structure with a relative clause that modifies it in the base. This second approach can be restricted within the limits imposed by Jackendoff on such rules by considering that \( \text{lo} \) in (17) is a nominalizing morpheme, providing rules such as (23a), to be compared with Jackendoff’s (23b):

\[
\text{(23) a. } N^i \rightarrow \text{lo-} A^i (\bar{S}) \\
\text{b. } X^i \rightarrow \text{Af-} V^i
\]

Such an analysis does not extent to (7), since \( \text{lo} \) is excluded from non-matching concealed questions (*\( \text{Sé lo con los problemas que se enfrentan} \)), even though it can very well precede an adverbial construction with an initial P in the exocentric relatives of type (17), as seen in the examples in (24) cited in (Alcoba 1981):

\[
\text{(24) a. Advirtieron } \text{lo en el acto que se presentaba el}
\text{They-noticed the in the act that himself showed-up the}
\text{bus-inspector}
\text{‘They noticed how suddenly the bus-inspector would show-up’}
\]

\[
\text{b. Criticaban } \text{lo } \{ \text{al final } \} \text{ que respondiste}
\text{at the end}
\text{They criticized the } \{ \text{at the beginning } \} \text{ that you-answered}
\text{‘They were critical of how late/how early you answered’}
\]
Since any P is possible in complements like those in (7) (i.e. the P in initial position is the one required by the embedded verb), it would be unreasonable to propose that P is a nominalizing category in Catalan and Spanish. In conclusion, (13c) is not a viable analysis for (7); the complements in those sentences are not NPs with a PP-head. In brief they are not exocentric relatives, irrespective of the correct analysis for (17).

Analysis (13d) is independently required for sentences (14), (15), and (16). It can also account for (7), given the treatment proposed for ordinary headed restrictives (i.e. deletion of a wh-phrase in COMP, or movement of an empty category to COMP). It is in accordance with X-syntax, but violates the subcategorization properties of the matrix verb at D- and S-structure, if base-generated, or at S-structure if the traditional raising analysis in some modality or other is adopted. Therefore (13d) is in conflict with the Projection Principle. Why is (13d) the less problematic proposal? The complement structures of (7) are possible in non-embedded contexts as well, as in (25) and (26). Obviously, in this situation there is no subcategorization problem and no conflict for the Projection Principle.

(25) a. En les dificultats {en qué} {que} ens vam posar!
   In the difficulties {in which} {that} ourselves we-have put
   'What difficulties we got ourselves into!'

   b. Amb la boca {amb que} {que} els esperen!
   With the mouth {with which} {that} them wait-for
   'With what a mouth they are waiting for them!'

(26) a. ¡Con los problemas {con que} {que} se enfrentan!
   With the problems {with which} {that} themselves they-face
   'What problems they face!'

   b. ¡De las tonterías {de que} {que} es capaz!
   Of the silly-things {that} {of which} he-is capable
   'What silly things he is capable of!'

The examples in (25) and (26) have exclamative interpretations exclusively, but this is a general characteristic of all concealed question/exclamative constructions when non-embedded:

(27) a. Sembla mentida la gente que hi havia!
   Seems lie the people that there was
   'It seems incredible how many people there were'
b. No sé la gent que hi havia  
Not I-know the people that there was  
‘I don’t know how many people there were’
c. La gent que hi havia!  
The people that there was  
‘How many people!’

As we can see in (27), the NP-complement can be exclamative or a question, depending on the nature of the matrix verb, but the unembedded structure in (27c) is only exclamative.

Also, the structures in (25) and (26) can be generated in non-subcategorized positions such as the TOP of the left-dislocation in (28):

(28) Jo en la forma que va començar allò no ho sé  
I in the form that started that-thing not it I-know  
(Calveras 1929:218)  
‘Me, in which way that thing started, I don’t know (it)’

Spanish is parallel: Yo, en la forma que comenzó eso, no lo sé. If we assume that the clitic ho/lo ‘it’ satisfies the subcategorization requirements of saber ‘to know’, then the Projection Principle is not at stake in this second case, and subcategorization is not the issue.

In view of all this, we conclude that the structures in (25) and (26) are PP-relatives, exactly as the complements in (13d). As we show in (Hirschbühler and Rivero 1981a), PP-relatives are independently required in the grammar of Catalan for Time/Place complements, and in more general terms, when a PP-antecedent is non-distinct from the phrase which is relativized. In an example such as (29a) there is a PP antecedent for the restrictive clause, and the complementizer que in COMP related to the (recoverable) deletion of a non-distinct PP phrase in that position. Alternatively, movement of an empty category to COMP dispenses with the need for a deletion rule.

(29) a. Això va passar en aquella ciutat que hi ha tants  
This happened in that city that there are so-many  
carrers per urbanitzar  
streets to plan  
‘This happened in that city in which there are so many streets to plan’
b. [ [en aquella ciutat] [ [qué] [hi ha.....] ] ]  
PP  
[S COMP S]
Spanish is similar. An example such as (30a) should therefore be analysed with a PP-relative complement.

(30) a. Tira al blanco que tiro yo
   Shoot at-the target that shoot (at) I
   'Shoot at the target I shoot at'

   b. [ [ al blanco] [ que tiro yo ] ]
   PP PP  S

That is to say, non-matching concealed questions/exclamatives are not the only structures requiring the PP-relatives generated in the phrase-structure rules of Catalan and Spanish.

In conclusion, the S-structure of the complements in (7), (9a), (10a), (14), (15), (16), (22b) are best analysed as PPs, along the lines of (13d). Given that the traditional proposal that moves the Preposition from the COMP to a pre-antecedent position is not viable within the EST, we must conclude that (13d) provides the D-Structure configuration. The same conclusion is reached on the basis of (11) and (15), which would require copying of the preposition rather than raising, and (16), requiring a different preposition to be introduced. Analysis (13d), which appears preferable on the different grounds we have discussed violates the Projection Principle because the subcategorization requirements of the matrix verb are not satisfied at S- and D-structure in relation to categorial information. However, as these complements are interpreted as questions or exclamatives, we propose that at the level of LF they have the form of questions or exclamatives, that is, they are Ss. Since the matrix verbs under discussion select S-complements at LF (in the sense of the Projection Principle), the sentences in (7) and related constructions have a well-formed LF, even though their D- and S-structures are problematic. Well-formedness at the level of LF is the relevant factor for their grammaticality. The rule that maps certain PP-surface structures into question/or exclamative S-structures at LF is language specific, but must be postulated, independently of (7), for non-embedded structures such as (25) and (26) where the Projection Principle plays no role. It is the existence of this language-specific rule that makes (7) viable in view of the Projection Principle. Under this proposal, the matrix verb in (7) assigns no θ-role to the PP itself, but only to the concealed question structure (i.e. to the S at LF that corresponds to the PP S surface structure). Thus the PP head of the relative plays no role with respect to the matrix verb, as expected from the lexical properties of that element.

Once we reach such a conclusion, cases like (31) are not surprising. If a verb subcategorizes for PP or for S as an indirect question, but not for NP, it may nevertheless take an NP-like relative as a concealed question.
Again, there may be a violation of categorial properties at D- and S-structure which is licensed by the well-formed LF. Verbs which exhibit those properties are recordar-se (not the transitive recordar) in Catalan, and acordarse 'to remember' in Spanish. We provide Spanish examples only but this argument was suggested by the discussion of the Catalan No m'en recordo dels llibres que hi havia 'I do not remember which books there were' vs. No m'en recordo dels llibres que hi havia 'I do not remember the books there were' in (Calveras 1929:216ff.).

(31) a. Acuérdate del (= de el) libro  
   Remember of-the  
   book  
   'Remember the book'

b. Acuérdate (de) que tiene el libro  
   Remember (of) that he-has the book

c. Acuérdate (de) qué libro tiene  
   Remember (of) which book he-has

d. *Acuérdate el libro  

e. Acuérdate los libros que compraron  
   Remember the books that they bought

Acordarse 'to remember' subcategorizes/selects PP complements with the P de as in (31a), P S-complements with the preposition de (the S can be a declarative or an indirect question) or S-complements lacking the initial de (also declarative or interrogative), as in (31c) but it does not allow for an NP-complement, as indicated by the ungrammaticality of (31d). However, example (31e) is grammatical with an indirect question interpretation: 'Remember how many books/which books they bought'.

The explanation is similar to the one provided for (7) and similar sentences. Example (31e) contains an NP-complement modified by a restrictive relative at D- and S-structure. This is in conflict with the Projection Principle, because the categorial properties of acordarse are not met at those two levels, As a concealed question the complement in (32) is treated as a S in LF, and at the level of representation the structure is well-formed, because acordarse subcategorizes/selects S as an indirect question. Notice that there is no P-raising analysis along the traditional lines of (11b-12b) that could be postulated in this case, given that there is no preposition.

If we assume that language specific mechanisms should be restrictive in their properties, we can consider the rule needed for non-matching concealed questions as the result of a special use of general mechanisms in the following way:
a) A PP after a transitive verb of knowledge/perception violates selectional requirements and leads to ungrammaticality, unless the phrase can be reanalyzed in a way compatible with those requirements. The same applies to the NP that appears to be the complement of the intransitive recordar-se/acordarse.

b) In a structure of the form $\chi P_1[\chi P_2 - \tilde{S}]$, $\chi P_2$ occupies a non-argument position, even in the unexceptional case in which $\chi P_1$ occupies an argument position.

c) The $\chi P_2$ in non-argument position can be i) an operator (i.e. a quantifier phrase or the Question/Exclamativ operator), or ii) the subject of a predication.

d) If $\chi P_2$ is taken to be the subject of a predication, $\chi P_1$ will be treated as an argument position of the $\chi P$ type (the usual relative). If, on the other hand, $\chi P_2$ is a non-argument position of the operator type, then the construction is considered a structure whose head is an operator (i.e. $\tilde{S}$ in our terms) at the level of LF. Thus the language specific device which is required for non-matching questions extends the functional interpretation of non-argument positions in a slight way.

Once the SS antecedent is considered to be in a LF operator position, it appears that the wh-operator for questions and exclamatives is the appropriate one for the interpretation. We stipulate this choice in (32) Notice that it could be the case that the specific character of the operator could be contextually defined. For example, after an indirect question type verb, only the question-operator interpretation is available. Also, in a non-embedded context, a construction like ¡Los niños que hay! is an exclamative, but its meaning correlates with a wh-exclamative, as in How many children there are!, or with an exclamative containing a non-wh quantifier, as in ‘There are so many children’! In conclusion, structures of the form of (32a) at SS are mapped into (32b) at LF; ?XP represents the question/exclamative operator:

(32) a. $\chi P[\chi P_i \tilde{S} [\text{COMP} \ [\text{wh-phrase}_i] \ S [\ldots t_i \ldots] ] ]$

b. $\tilde{S} [\ ?XP_i \tilde{S} [\text{COMP} \ [\text{wh-phrase}_i] \ S [\ldots t_i \ldots] ] ]$

To summarize, if standard relative clauses are seen as predicking a property of the antecedent, then they are one-place predicates, that is, open sentences with one open position. The open-sentence status can be attributed to the relative $\tilde{S}$ in our non-matching concealed questions as well. However, at the level of LF, the “antecedent” in this last case is treated as an operator with respect to the open sentence represented by the embedded clause. Intuitively, non-matching concealed questions are relative clauses at
D- and S-structure, and questions at LF.

The result of (32) is an adjunction structure with the dominating node satisfying the lexical requirements of the matrix V.3 In non-embedded contexts such as (25) and (26), the effect of (32) is identical, but there are no lexical requirements that the structure must satisfy. Throughout the discussion, we take S to identify a syntactic category at LF, and not a semantic type like Q(uestion) or E(xclamative), since such a distinction is derivative in non-matching constructions, and plays no role with respect to the issues discussed in this paper.

Remember that we have syntactic reasons not to generate the first PP in (7a) in COMP along the lines of (13a), and that we have opted for an antecedent-type analysis as in (13d) for the various reasons presented. Even if a COMP-analysis had been postulated, the same interpretive language particular mechanism would be required. Namely, in (14), the relative phrase in COMP and the embedded S would need to be treated as the open sentence, and the antecedent-like constituent located in the COMP would have to count as a question-operator in the LF component.

The Projection Principle restricts the possible mappings between the three syntactic levels of representation. In Chomsky’s conception, categorical and structural identity characterizes this mapping. Thus, lexical properties connected with θ-role assignment and subcategorization are structurally and categorially encoded in an identical manner at three levels of analysis (i.e. the relationship between DS, SS, and LF is one of identity). This conception appears clearly in the discussion of of-insertion (Chomsky 1981:pp. 51ff.) in such examples as destruction of the city (i.e. (2i) in the quote below):

At D-structure, destruction (or, perhaps, its head destroy) subcategorizes and θ-marks the NP object the city in (2i). If the projection principle is valid, this must also be true of S-structure and LF. There is no problem if the of-insertion rule is an adjunction rule forming the NP [NP of NP] as in one of the options we have just considered, or if the of-phrase is base-generated as PP (cf. note 21). If the of-insertion rule creates a PP, however, we must continue to hold that destruction subcategorizes its NP object at S-structure and LF. Note that still another possibility would be to assume that the of-insertion rule forms a neutralized NP-PP of the form [-V], in which case, again, no problem arises and this category will share properties of both NP and PP.

Under this approach, rules of correspondence that translate one category into another when going from one level to the other, or structure-building rules of the type found in “On binding” violate the Projection Principle.

A different conception of the correspondence between the different levels of analysis does not require identity but establishes for each level
the relevant lexical requirements to be satisfied and their not necessarily identical counterparts at the other levels (Marantz 1981).

The constructions discussed in this section are problematic for either conception of the Projection Principle (i.e. for the satisfaction of subcategorization properties at the various syntactic levels). If the relationship between SS and LF is one of categorial identity, the constructions in (7) raise two issues:

a) Assuming that all types of questions are treated in a unitary way as \( \hat{S} \) at the level of LF, then the requirement of categorial identity between SS and LF is not met by our constructions, nor by ordinary concealed questions of the NP-type, such as (27b). Under the identity-interpretation of the Projection Principle, it is perhaps better to consider the ordinary type of concealed question as an NP both at S-structure and LF, providing the necessary semantics for NPs of different types. For this approach within a Montague-framework see Heim (1979).

b) There is a fundamental difference between the constructions called concealed questions in the literature, which appear to be widespread among languages, and non-matching concealed questions in Catalan and Spanish. The verbs under consideration do not select PP-complements (i.e. no PP without a relative clause is grammatical, as seen in (6)). Alternatively, they do not select an NP in cases like (31). So non-matching concealed questions constitute definite subcategorization violations at S-structure, while the ordinary concealed question does not.

Turning to the conception of the Projection Principle not based on identity but on the notion of counterpart, correspondence rules could be set up to relate NPs satisfying the surface requirements of transitive verbs and the category Q (our \( \hat{S} \)) at LF, much in the spirit of the proposals found in Grimshaw (1979). In other words the S-structure of *Saba la hora que es* ‘He knows the time it is’ would be well-formed and would have as a counterpart a LF with a Q-complement, posing no problems for this view of the Projection Principle.

However, our non-matching concealed questions would still fall outside of this theory, given that at S-structure they violate subcategorization requirements. Whatever version of the correspondence between S-structure and LF is adopted, the non-matching concealed questions remain as constructions that do not fit the Projection Principle if seen as encoding categorial information at D- and S-structure. In the case we have discussed, the properties of verbs of knowledge and perception are satisfied at LF. As Chomsky points out (1981) lexical properties must be satisfied at this level by definition. We have seen that deviations from subcategorization/selection at D- and S-structures are permitted if the constructions involved are mapped into a well-formed LF structure. Thus, a set of verbs and a preposition can take as S- and D-structure complements any structure that
is mapped into a question or an exclamative. All of this raises the issue of the parametrization of the Projection Principle, or a reconsideration of the theory of selection behind (1). We will address this topic in section 4.

3. FREE RELATIVES AS CONCEALED NON-MATCHING QUESTIONS AND THE PROJECTION PRINCIPLE

With the proposals of the previous section in mind, we turn to the sentences in (8). For reasons to be given later, we consider that del que in (8a) (which is the contracted form for either de el que or de lo que in different varieties of Catalan) and Spanish de lo que in (8b) contains a relative form, not a question word. Thus, the complements in (8) are free relatives whose initial constituent is a PP, and they occupy the complement position of a verb that subcategorizes for NP or S, but not PP.

Before we discuss the relationship between (8) and (7) on the one hand, and the Projection Principle on the other in 3.2, we summarize the analysis required for free relatives in Catalan, and apply it to Spanish in 3.1.

3.1. The analysis of free relatives

In (Hirschbühler and Rivero 1981a-b, 1983), we distinguish two types of Catalan constructions that can be termed free relatives, as in (33):

(33) a. Qualsevol que digui això ment
    Whoever that would say that lies
b. Qui diu això ment
    Who says that lies

Spanish examples (34a-b) are parallel to (33a-b) respectively, and the Catalan analysis extends to Spanish, as we shall see.

(34) a. Cualquiera que diga eso miente
    Whoever says that lies
b. Quien dice eso miente

Examples (33a) has qualsevol ‘whoever’ as antecedent or head of the relative, and the que form in COMP, as in (35a). Example (33b) has the relative qui ‘who’ in COMP, and the antecedent is an empty category, as in (35b) (we leave open the nature of such a category). Traces are omitted.

(35) a. [ [ Qualsevol ] [ [ que ] [ digui això ] ] ]
    NP NP S COMP S
b. [ [ e ] [ [ qui ] [ diu això ] ] ]
    NP NP S COMP S
We treat the antecedent as a maximal projection NP, but nothing in our analysis hinges on this.

The main argument for the structures in (35) relies on so-called "matching" properties (Grimshaw 1979). Free relatives of type (35a) must be matching both in subcategorized and non-subcategorized positions. That is, the category of the initial phrase, and the first phrasal constituent that dominates the relative construction must be of the same type. In (33a), *qualsevol* "whoever" is an NP inside of an NP ([NP [qualsevol] ...]), and the construction is "matching". Consider (36a-b); in (36a) *invitar* requires an NP complement. *Qualsevol* 'whoever' is external to COMP. Two structures can be assigned to (36a): *invita NP[ PP[amb qualsevol] ...], and invita NP[ PP[amb qualsevol] ...]. The first one does not satisfy the requirements of the verb, and the second one violates \(\bar{X}\)-requirements. In (36b), the topic position contains a PP, and correlates with an NP in the following clause. The deviance is due to the same factors that rule out examples such as With John, I know him in English; the topic that corresponds to an NP denoting an individual must be an expression capable of denoting an individual, and PPs are not.

(36) a. *Invita amb qualsevol que t'en anirás
    Invite with whoever that you-will-leave

    b. *Amb qualsevol que t'en anirás, l'invitarem
       With whoever that you-will-leave, him we-will invite

Examples (37a-b) show the same for Spanish:

(37) a. *Invita con cualquiera que te vayas

    b. *Con cualquiera que te vayas, lo invitaremos

The matching properties of relatives of type (33b) are different. This construction must be matching in subcategorized position, as in (38a), but need not be in non-subcategorized position, as indicated by the grammaticality of (38b):

(38) a. *Invito amb qui t'en anirás
    I-invite with whom you-will leave

    b. A qui li toque el bon mos, que roseque el os
       To whom him falls the good bit, that he-chew the bone
       'Let the one who gets the good bite chew the bone'

      (Calveras 1930:177)

In (38a) the structure is *[NP [pp amb qui] ...]. In (38b) we assume
We refer the reader to the references given for detailed discussion. The subject position exhibits parallel non-matching properties.

The phrase in the Top node in (38b) is an NP, and the initial constituent with lexical material is a PP. This PP can only be in COMP, since a PP cannot be the head of an NP, given the conventions of X-theory. Assuming that in both (38b) and (38b) [amb qui] and [a qui] occupy the COMP position, the contrast in grammaticality follows from subcategorization and accessibility to COMP in free relatives. Why is (38a) ungrammatical? When the antecedent is lexically empty, the material in COMP must satisfy the subcategorization requirements of the matrix verb (Groos and van Riemsdijk 1981). This predicts that free relatives of type (33b) must be matching in subcategorized position. (Invito [ [ e [ [ [ qui] NP NP \tilde{S} COMP NP [has invitat] ] ] ], excluding (36a), but can be non-matching in non-subcategorized positions, allowing (36b).

Spanish offers a slightly more complex situation. Examples (37a-b) show that cualquiera is in antecedent position; the ungrammaticality is due to the same factors that exclude (36a-b) in Catalan. Examples (39a-b) are parallel to (38a-b) in structure and meaning:

(39) a. *Invito con quien te irás
    b. A quien le toque el buen bocado, que roa el hueso

Example (39b) must have a quien ‘to who’ in COMP for the reasons given for (38b) in connection with \( \tilde{X} \)-conventions. Under the assumption that the free relative occupies an NP-node in TOP given its interpretation, a non-matching free relative is allowed in non-subcategorized position, as in Catalan. Example (39a) is more complicated.

Human objects are preceded by a, so that we have Invito a María ‘I invite Mary’ and not *Invito María. Under such a situation, it could be proposed that invito subcategorizes for a PP, not an NP, human complement. Then strict subcategorization alone could not account for the deviance of (39a), since we have the appropriate structure for matching purposes: \([VP \text{ invito } [PP [PP con quien ] ...]]\). However, it has generally been assumed that direct objects of the a + NP form are NPs not PPs (see Jaeggli 1981 for a recent argument to this effect). Under this assumption, invito subcategorizes for NP in all cases, and the ungrammaticality of (39a) is due to the presence of a non-matching free relative in a subcategorized position: \([ \text{ invito } [ [ [ con quien ] ... ] ] \). A similar approach can be used for the varieties of Catalan where Invito a qui has invitat alternates
with *Invito qui has invitat* 'I invite who you have invited'. Actually, if the phrase in COMP satisfies the requirements imposed by the verb on its complements, then both subcategorization and selectional properties become relevant. The preposition *amb/con* 'with' is not subcategorized/selected by *invitar* so that (39a) is deviant even if one considers that *invitar* subcategorizes for a PP with *a*.

In conclusion, Catalan and Spanish place the relative phrase in (33b) and (34b) in COMP, a position which is accessible to the subcategorization requirements of the matrix verb.

### 3.2. El que/lo que free relatives and the Projection Principle

Because of the morphological characteristics of *el que/lo que* sequences, the free relatives that contain this string can in principle be analysed as headed constructions with *el/lo* in antecedent position and the complementizer *que* in COMP, equivalent to the *qualsevol* relatives in (33a) with *el/lo que* as a compound relative in COMP (as the *qui* relatives in (33b)). Independently of the analysis which is selected, the constructions in (8) are problematic for subcategorization.

If the complements in (8) are NPs subcategorized by *observar* and *saber*, with an empty antecedent and a relative with *el que/lo que* in COMP, they contain a non-matching free relative in subcategorized position. This situation should make the sentences ungrammatical, because the COMP is accessible to the subcategorization requirements of the matrix verb, and the PP in COMP cannot satisfy them.

\[(40) \quad \text{a.} \quad [\text{observa} [ [e] [ [\text{en} \{el\} \text{que}]])
\]
\[\text{VP} \quad \text{NP} \quad \text{NP} \quad \tilde{S} \quad \text{COMP} \quad \text{PP}
\]
\[\text{[ens vam posar]} ] ]]
\[S
\]
\[\text{b.} \quad [\text{se} [ [e] [ [\text{en} \{lo\} \text{que}])][\text{se enfrentan}]]]]
\[\text{VP} \quad \text{NP} \quad \text{NP} \quad \tilde{S} \quad \text{COMP} \quad \text{PP}
\]
\[S
\]

It could also be proposed that these free relatives can be "matching", in the sense that their antecedent contains a null-PP (given the existence of phrase-structure rules of PP-relatives), as in (41):

\[(41) \quad \text{a.} \quad [\text{observa} [ [e] [ [\text{en} \{el\} \text{que}]])
\]
\[\text{VP} \quad \text{PP} \quad \text{PP} \quad \tilde{S} \quad \text{COMP} \quad \text{PP}
\]
\[\text{[ens vam posar]} ] ]]
\[S
\]
\[\text{b.} \quad [\text{se} [ [e] [ [\text{en} \{lo\} \text{que}])][\text{se enfrentan}]]]]
\[\text{VP} \quad \text{PP} \quad \text{PP} \quad \tilde{S} \quad \text{COMP} \quad \text{PP}
\]
\[S\]
But if the lexical material in COMP must satisfy the subcategorization requirements of the matrix when the head of the relative is null, the situation is the same as for (40).

Under the analysis that places *el/lo* in antecedent position, and *que* in COMP, we face the same problems we encountered in the discussion of (13c) and (13d). If *en el/lo* and *con lo* are PPs dominated by NPs, the requirements of the $X$-theory are violated:

\[(42)\]
\[
\begin{array}{ll}
\text{a. } & \left[ \left[ \text{en } \{ el \} \right] \left[ \text{que ens vam posar} \right] \right] \\
\text{NP PP} & S \\
\text{b. } & \left[ \left[ \text{con lo} \right] \left[ \text{que se enfrontan} \right] \right] \\
\text{NP PP} & S \\
\end{array}
\]

If these constructions are PP-relatives, they violate subcategorization requirements at S-structure, and D-structure in the base-analysis we have adopted before:

\[(43)\]
\[
\begin{array}{ll}
\text{a. } & \left[ \left[ \text{en } \{ el \} \right] \left[ \text{que ens vam posar} \right] \right] \\
\text{PP PP} & S \\
\text{b. } & \left[ \left[ \text{con lo} \right] \left[ \text{que se enfrontan} \right] \right] \\
\text{PP PP} & S \\
\end{array}
\]

In brief, the question raised by the structures in (8) are exactly those raised by (7). Analyses (40), coupled with Accessibility to COMP, and (41), (43) provide violations of subcategorization requirements, much like (13d) did. If (40), (41), and (43) violate subcategorization at S-structure, the language-specific rule independently required for the non-matching concealed questions/exclamatives of the previous section can treat the free relatives that complement *observar* and *saber* as $\tilde{S}$s at the level of LF, allowing the conflict with the Projection Principle discussed in the previous section. In other words, whether accessibility to COMP is required for free relatives in analyses like (40) or (41), with *lo que* as a relative in COMP, or whether *lo que* is treated as a non-constituent with *lo* in antecedent position and *que* in COMP as in (43), the structures in (8) can receive the same treatment as those in (7).

The argument developed for (31) in relation to *acordarse* ‘to remember’ is also possible here. *Acuérdate lo que dijeron* is grammatical and interpreted as an indirect question (‘Remember what they said’), even though *acordarse* does not subcategorize/select an NP.

There are two possible analyses available for (8) which comply with the Projection Principle as interpreted so far. In the first one, (8) is an ordinary indirect question with a PP in COMP containing the question word *lo-que/el-que*. In the second one, the complements of (8) are declarative-like $\tilde{S}$s with *lo-que* in COMP.
First, if *lo-que/el-que* forms are question words, not relative phrases, then (8) would not provide a violation of the Projection Principle under the usual analysis for indirect questions:

(44) a. [ observa [ [ en lo que ] [ ens vam posar ] ] ]
    VP $\bar{S}$ COMP $S$
    
    b. [ sé [ [ con lo que ] [ se enfrentan ] ] ]
    VP $\bar{S}$ COMP $\bar{S}$

An indirect question is an $\bar{S}$ with a content in COMP which is not accessible to the matrix for categorial purposes at S-structure or at LF (the category in COMP is immaterial as long as it is a question phrase). At each level of representation, the $\bar{S}$, not the question-word in COMP, satisfies the requirements of the matrix. Although the examples in (8) are interpreted as indirect questions, it is rather clear that they are not syntactic questions, because *el que/lo que* forms do not have the distributional properties of question words (i.e. they do not belong to the interrogative paradigm, only to the relative one).

1) The first argument to consider *lo que/el que* a relative sequence is quite clear in Spanish, but subtle in Catalan. The form of the argument was nevertheless suggested by the discussion of *el què* in direct questions in Oriental Catalan found in Calveras (1929:189, 221). Let us first consider Spanish.

As Plann (1980) points out, unproblematic interrogative words can appear in direct questions, but *lo que* cannot. Compare *cual* in (45a) which can only be used in questions, not in relatives, and *lo que* in (45b):

(45) a. ¿Cuál prefieres? ‘Which (one) do you prefer?’
    Which you prefer
    b.*¿Lo que prefieres?

For an *echo question*, Spanish has (46), but the article form differs from that of (45b) (*el* is masculine singular, *lo* is neuter), and *què* is the stressed interrogative (*que* in (45b) is unstressed):

(46) ¿El qué prefieres? ‘So, what (did you say) you prefer?’
    The which you prefer

When the question word remains *in situ*, the same grammaticality contrasts obtain:

(47) a. ¿Has dicho qué? ‘You said what?’
    you-have said what?
b. *¿Has dicho lo que?

c. ¿Has dicho el qué? ‘You said what?’

you-have said the what

Example (47c) is an echo question exclusively. From these two patterns, it can be concluded that *lo que*-forms are excluded from direct questions in Spanish.

In Catalan, the situation is much more complex. First, the Oriental variety has an *el*-article both for the masculine singular and the neuter; the Occidental variety has *lo* for both. In both varieties, the equivalent distinction to Spanish (45c) vs. (46), or (47a) vs. (47c) does not rest on the article forms, which are homophonous, but on the stressed *qué* vs. the unstressed *que*, which differ in vowel quality. Second, *qué* is both a relative word, and an interrogative one, while Spanish *qué* (as in 47c)) is interrogative exclusively. The patterns of Catalan in this area are interpretable in many more ways than the equivalent situation in Spanish. Checking the properties of *lo que* forms in Occidental Catalan, we found that (48a-b) are grammatical and interpreted as echo questions.

(48)  a. Lo qué has dit? ‘You said what?’
             The what you-have said
    b. Has dit lo qué? You-have said the what

We assume that *Lo que has dit?* and *Has dit lo que?* are ungrammatical. The difference between these last sentences and (48) is difficult to hear, but it is always possible to emphasize the *qué* form in (48), which is not a property of *que*. For Oriental Catalan, Calveras gives *El qué diu?* as well-formed (*el què'val en el llenguatge popular per què, interrogativ de cosa (1929:189)), and omits discussion of *el que* in this context. We assume he did not find it used in direct questions, or judged it impossible. His omission provides indirect support for the conclusion that *el que/lo que* forms are not interrogative sequences.

Both direct and indirect questions may contain several interrogative wh-phrases in multiple questions, but *lo que/el que* is excluded in situ from these constructions, as in (49c) and (50c):

(49)  a. Quina noya ha parlat de qué? Which girl has-spoken of what?
    b. T’estic preguntant quina noya ha parlat de qué
       you I'am asking which girl has spoken of what
       ‘I am asking you which girl spoke of what’
    c. *T’estic preguntant quina noya ha parlat del que
(50)  
  a. ¿Quién hace qué?
      'Who does what?'
  b. Me pregunto quién hace qué
      'I wonder who does what'
  c. *Me pregunto quién hace lo que

This is expected if el que/lo que forms are not question words but relatives, and as a result restricted to COMP.

3) In Spanish structures that Plann (1982) labels reported questions, the que complementizer may precede any type of question word, as in (51a-b), but lo que is excluded in that environment:

(51)  
  a. Pregunto que con quién vienen
      I-ask that with whom they-arrive
      'I ask who they arrive with'
  b. Pregunto que por qué lo hacen
      I-ask that why it they-did
      'I ask why they did it'
  c. *Pregunto que lo que hacen

Plan proposes to analyse (51a) as in (52).

(52)  

Only those structures which are well-formed questions can be embedded under $S_2$. Under this analysis, the ungrammaticality of (51c) reduces to the deviance of (47b), with the conclusion that lo que is not a question word.

In brief, if the el que/lo que forms are question words, then their distri-
bution has unexpected restrictions: a) they only appear in embedded contexts; b) they only appear in COMP; c) they have to be excluded from the lower COMP of the reported question construction in (5). These are all restrictions applying to relative words, so we conclude that these forms are relatives, and that they can be the initial constituent of ordinary free relatives.

Given a) the lexical properties of observar and similar verbs in (8), b) the structural characterization of free relatives with null antecedents discussed for (33b), and c) the properties of Wh-movement, structures such as (40) must be generated in Catalan and Spanish, if lo que and el que forms are relatives, as we just argued. Our proposals account for the fact that the free relatives in (8) are not interpreted as the usual relatives (i.e. as a referential expression or a name, as a generic-like NP...). As already pointed out, the free relative in (40) occupies a subcategorized position. Under the assumption that accessibility to COMP comes into play in such a case, (40) is filtered out at the level of S-structure (i.e. after Wh-movement) because the lexical properties of the matrix verb are not satisfied by the element in COMP. In this sense, (8) is parallel to such ungrammatical sentences as (28a) and (39a) (under analysis (40)). So (8), (38a) and (39a) are filtered out as ill-formed as “regular” free relatives (i.e. as constructions which fall within the class of referential-like expressions). However, there is an important lexical difference between the observar/saber verb in (8) and invitar in (38a-39a). While invitar cannot have a concealed question as complement, observar and saber can. Thus, the solution developed in section 2 for (7) is available for (8), but not for (38a-39a). In conclusion, the complement structures in (8) are non-matching concealed questions which can be analysed as non-matching free relatives at S-structure (i.e. [ [ e ] [ [ PP] [ ....] ] ] ) and Šs at the level of LF. NP NP Š COMP S

They are not indirect questions with a question word in COMP at S-structure (an analysis which leaves unaffected the devices proposed for (7) in the previous section in any event).

Let us now turn to the second analysis that treats (8) in accordance with the Projection Principle. Given the subcategorization properties of verbs of knowledge, perception and saying, the declarative-Š frame allows (8) to be analysed as in (53), proposed by Plann (1980) for Spanish, with incidental differences for the purposes at hand. Under the analysis, (8) does not contain a free relative of type (35b).
Tree (53) satisfies the Projection Principle because its complement meets the subcategorization frame of the matrix at all levels: D- and S-structure, and LF. Since we have concluded together with Plann that lo que is not a question word, we assume that an additional interpretative device is required to allow for the indirect question interpretation. Perhaps it is a language-specific rule giving this S a translation identical to that of a similar clause containing an interrogative wh-phrase in COMP, which may be equivalent to treating this $S$ as a concealed question with no question-word. As we have shown, in a grammar that accounts for those structures in (7), the properties of (8) follow, assuming the proposed analysis for free relatives. Structure (53) is not problematic for our proposals. However, its postulation, and the necessary interpretive rule, achieve the same results we have achieved for (8) assuming only one type of structure for all free relatives. For the various reasons given in the previous section, (53) cannot be extended to cover (7).\(^5\)

In brief, we conclude that the free relatives in (8) are a subset of the non-matching concealed questions and conflict with the Projection Principle at S-structure, but are well formed at LF.

4. THE PARAMETRIZATION OF THE PROJECTION PRINCIPLE\(^6\)

In section 2, we saw that an interpretation of the Projection Principle which requires categorial identity at the three levels of representation is in conflict with the properties of non-matching concealed questions. The same is true of an interpretation requiring well-formedness at each level without strict identity between the three of them. In the last section, we showed that the subset of non-matching free relatives which are interpreted as questions raise similar problems. In brief, the level of LF has priority over the other levels of representation for the satisfaction of lexical requirements; D- and S-structure "requirements" are simply that the argument(s) of the verb be of a type that can be mapped into the appropriate category at LF.
In view of all this, we can reexamine the Projection Principle from two different and logically independent perspectives.

First, is the status of the three levels of representation in part (iii) of the Principle the same? If lexical requirements are not met at all three levels in a given case, which level is essential for the satisfaction of selection? We have already pointed out that selection must be satisfied at LF by definition (Chomsky 1981), so this level has priority over the other two.

Second, what is involved in the term selection in (1)? Should the different parameters behind this theory enjoy the same status within the formulation of the principle? Selection involves information about several parameters that may be logically independent of each other. Concentrating on direct θ-marking exclusively, we should distinguish the following:

I. The number of arguments a lexical head requires (i.e. the A-positions to be filled)

II. The semantic value of each argument, including
   a) the θ-role the argument bears (Agent, Theme, etc.),
   b) the semantic type of the argument (Question, Proposition, Referential Expression, etc.)

III. The categorial encoding of each argument (NP, S, etc.)

As we have already mentioned, one approach to the Projection Principle (Chomsky 1981:51ff.) is to assume that requirements of types I, II, and III be exactly the same at all three levels. Let us label this interpretation the strong version of the Principle.

A weaker interpretation that we will adopt is proposed by Pesetsky (1982:17). He requires that the number of arguments (information of type I above), and the θ-positions (information of type II) remain the same at all three levels, but makes the principle blind to categorial properties (information of type III):

(54) Representations at each syntactic level (LF, D-structure, S-structure) are projected from the lexicon, in that they observe the θ-marking properties of lexical items.

Categorial properties are not projected to every level of representation, but must be satisfied at LF.

Lexical requirements of type I and II are the parameters of the theory of selection that must fall under the Projection Principle. They enjoy a degree of universality that requirements of type III do not. Considering the two cases mentioned in this paper, notice that the appeal to the Projection Principle to rule out the Raising-into-Object analysis for (2a) deals with I and II, prohibiting in particular that the number of arguments of believe be different at each level, but does not seem to rely on III. Also,
the non-matching concealed questions comply with parameters I and II, bit not III; the Projection Principle as formulated in (54) predicts this to be the case.

However, the strong version reflects the observation that in most instances categorial identity between SS and LF is the norm. Under the weaker version adopted here, divergence between the categorial make-up of arguments at LF and SS, or identity between the two levels, appears to be equally unmarked and expected cases. Also, the situation captured by the strong version is the ideal one for learnability purposes, because it implies that the mapping between SS and LF is trivial. Nevertheless, under the assumption that in the unmarked situation, the theory of grammar allows no procedures of correspondence between levels of representation other than those ensured by Move α and indexing mechanisms, categorial identity between the levels follows without stipulation in most cases. The non-identity of categories will arise in situations where a particular mechanism establishes a correspondence between a category of one level, and a different one at another level, and such cases will count as marked. Thus, we can obtain the results of the strong version of the Projection Principle by adopting the weaker version excluding information of type III.

Under this approach, non-matching concealed questions are marked because they require a special use of an existing mechanism, as seen in the discussion of (32); but these constructions are not in conflict with a version of the Projection Principle that includes no stipulation about categorial identity. At the level of LF, non-matching questions comply with lexical requirements of all three types; at D- and S-structure they comply with requirements of type I and II, in accordance with the principle.

A topic for further research is the types of (marked) categorial divergences that are allowed, or, to express it in somewhat different terms, the types of marked mechanisms that can be considered learnable. The formulation in (32) suggests the postulation of a category-switching rule as the mechanism which is language-particular in our case. However, the marked aspect can also be seen as an extension of the phrase-structure projection mechanisms required for operator-type structures at LF. Under this view, there is no category-switching “rule” between SS and LF, or vice versa; rather, at the level of LF, the initial element in the structure of a non-matching question is considered an operator, which leads to the functional definition of the dominating category as $\bar{S}$.

The marked properties of the non-matching concealed questions show the priority of LF over other levels of representation for the satisfaction of lexical properties, and indicate that categorial requirements are expressed only at that level.
NOTES

1. The stylistic and normative status of these constructions is different in Spanish and Catalan. For Spanish, both (7b) and (8b) are discussed together as ‘incorrecciones toleradas’ in Moliner (1975; see relativo), and the second type is considered extremely frequent, but no different in status from the first. See fn. 4 as well. For Catalan, Calveras (1929:§30-34) offers a lengthy and very suggestive discussion of type (7a), but only incidental comments on type (8a) as a subcase (1930:p.182). Normative concerns are not usual in this grammarian, but here and there he considers that these constructions belong to the popular language. Very recently, in a generative treatment of relatives, Viaplan a (1980:194) – the source of (7a) – accepts the traditional analysis schematized in (11b), and considers this aspect of Catalan grammar very much alive, but offers no comment on their normative or stylistic ranking: ‘En el caso del català, aquesta tendència [to move the preposition before the antecedent] és extremament viva i pot afectar la majoria de preposicions’. The construction in (8b) is a vexing one for Catalan grammarians, who feel it is calqued from Spanish. According to Badia (1962, I:268), it is frequent but incorrect: ‘En una frase como “ignoro lo de que eres capaz”, al sentirse toda la oración sustantivada como término de la preposición, está se antepone: “ignoro de lo que eres capaz”, tipo más corriente; esta combinación “de lo que” provoca una traducción catalana incorrecta por castellanizante, aunque muy extendida: ignoro del que ets capaç [n.a.] que no se puede aceptar de ninguna manera’. In this paper, we show that the mechanisms required for (7a) are no different from those required for (8a): if a grammar generates (7a), it generates (8a). Thus, regardless of the sociolinguistic status of (8a), and irrespective of whether it is an instance of calquing, the grammar of Catalan contains independently required formal devices that lead to its generation. We address this issue in more detail in a paper to be published in the Serie Lingüística of the Universidad Autónoma de Barcelona.

2. The problems that (7) and similar structures raise for the Projection Principle are independent of the derivational history. In the traditional raising analysis, or in the base-analysis we propose, the S-structure representation is in violation of lexical properties. However, the traditional analysis translated into transformational terms requires a structure-building movement rule, and may violate Subjacency if NP and S are bounding nodes. In addition, copying rather than raising would be required for (14-15), and insertion of a different preposition, without raising, for (16). The raising analysis could not account for (31). For a formalization of the traditional analysis for (7b) in an early transformational framework see Cresse y (1970). Within an Aspects-model see Viaplana (1980:203) for Catalan.

3. A reviewer raises the question of how the NP in (31c) gets case; as acordarse does not subcategorize for NP, it should not assign objective case. We may instead ask whether the NP should be case-marked by the matrix V or bear case at all. At the level of LF, the NP in (31c) is roughly of the form $\mathbf{g}^1 \mathbf{op}_1 [\mathbf{los libros}] \mathbf{g} [\mathbf{que}, \mathbf{com-}$ $\mathbf{praron} \mathbf{t}_1]$ with los libros treated as an operator (i.e. a phrase in non-A position). Several possibilities are open as to case-marking. Since the phrase is in non-A position, it may escape the case filter, if such a device is reserved for NPs in A-position (Chomsky 1981, Safir 1982); thus, los libros could be caseless with no ill effects. Alternatively, it could be proposed that los libros inherits its case through the index it shares (after SS) with the trace in A-position in the embedded clause. Also, the NP could be case-marked by default by the mechanisms required for non-A positions such as the topic of Left-dislocations. The important factor seems to be that the relative construction does not count as a NP argument of the V, so case devices for non-A positions should be appropriate. Notice that the same issue arises for (27c).
4. We consider that *lo que* in COMP is a compound relative (for reasons given in Rivero 1980c); Plann treats the sequence as an article *lo* surfacing within a relative phrase, followed by the complementizer *que*. The internal analysis of *lo que* does not bear on the issues we have addressed.

5. Plann (1980) considers structures of type (7b) ungrammatical, and eliminates them completely from the analysis, because she finds speakers who reject them. If normativism is not at work (or perhaps stigmatization) when a speaker rejects (7b) (see fn. 2), and there is a genuine dialectal difference behind the ungrammaticality of (7b) and the well-formedness of (8b) in such a situation, then there could be a variety of Spanish with only analysis (53) for (8b), no violation of the Projection Principle, if it combines categorial and structural requirements and no structures like (13a) available.

6. We thank the reviewer who suggested that we consider the parametrization of the Projection Principle, and pointed out the connection between our proposals and Pesetsky's work, which became available during the reviewing process for the present paper.

REFERENCES


Calveras, J. 1929-30. 'La forma “que” del relatiu català', *Anuari de l'Oficina Romànica de Lingüística i Literatura*, 2.185-254. 3.177-243.


*Department of Linguistics*

*University of Ottawa, Ottawa, Ontario, Canada.*