

# The structure of Comp in Slavic: some evidence from Slovak

LUIS A. SAEZ

(Universidad Autónoma de Madrid - M.I.T.)

## Abstract\*

*Several phenomena concerning multiple questions suggest that Slavic languages can be split in two groups. Rudin (1988) under the GB-framework, distinguishes those groups by means of a condition saying that multiple adjunction to the Spec of CP is/is not possible. We point out that several problems arise with that view, and propose that it is the (un)availability of Comp recursivity that discriminates such languages. It is the availability of Comp recursivity that accounts for the following facts: a) each preposed wh-element finds a landing site after movement (when no Comp recursivity exists, multiple wh-preposing will be understood as actually involving movement of a verbal projection to the Spec of CP); b) multiple wh-movement from inside an embedded clause is not possible (we give a Subjacency-based account once CP has been adopted as weak barrier); c) clitics have to appear to the right of the first wh-element. c) follows from further proposals about T in Slavic languages: a) it is generated above Agr; b) it is the landing site for clitics; c) it has to raise to C at SS, since it is an operator. The main burden of proof comes from Slovak, taken as paradigm of the Slavic group exhibiting Comp recursivity.*

## 1. Some data: Slovak in the context of Slavic languages.

Slovak multiple questions share with other Slavic languages the property that all the wh-elements obligatorily undergo movement at SS to the left margin of the sentence. This is not a property common to all languages. For instance, let us take an English multiple interrogative such as (1):

(1) Who did what?

In this sentence, there are two interrogative wh-elements, *who* and *what*. However, one of them (*what*) has to stay in situ. It is not possible to move it to the left, as (2) shows:

(2) \* Who what did?

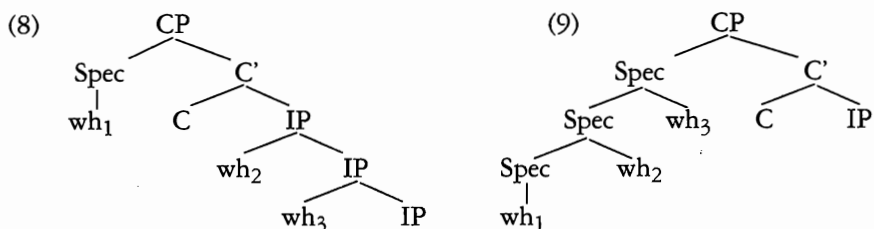
\* This paper has been supported by a grant given by the Socialist Republic of Czechoslovakia. It has also been partially supported by the Spanish Ministry of Foreign Affairs. Thanks are due to J. Fastorová, J. Furičková, S. Kuchtová and M. Puchovsky for providing me with the Slovak data, as well as to the audience of the first Congress "Datorren Sintaxiazko Hitzaldiak" (Euskal Herriko Unibertsitatea; may-june 1990, Gasteiz, Basque Country, organized by K. Sainz), where this paper has been presented. Finally, thanks are also due to an anonymous reviewer or ASJU for his/her comments to a first version of this article. Of course, I am the only responsible for all the possible shortcomings which still remain.



In the next sections, we will try to provide some hypothesis concerning the properties of Comp in Slavic languages, which account for these data. Our main concern will be Slovak, a language which, to our knowledge, has scarcely been studied within the Government and Binding framework.

## 2. Rudin's (1988) account.

Rudin, using the Subjacency principle, which is parametrizable, relates the possibility of multiple extraction from inside an embedded clause with the Comp properties of the different Slavic languages. This is performed in the following way. First, she proposes that the underlying Comp structure of sentences in (4) is not the same for all the cases. Indeed, for Cz, S-C and P, such a structure would be as in (8), whereas for Bulgarian it would be as in (9):



She refers to the first group (Cz, S-C and P) as (-MFS) languages ("Non Multiply-filled Spec-CP Languages"): languages where the Spec of CP cannot be simultaneously filled by two or more wh-elements. By contrast, Bulgarian would be a [+MFS] language: as shown in (9), this means that different wh-elements can be adjoined to the Spec of CP.

The next step of Rudin's account has to do with the Bounding Theory: a wh-element which, along its movement, finds a Spec of CP already filled by other wh-element, will only be able to cross that CP if it is adjoined to such a Spec. Otherwise, Subjacency is violated. Since in (-MFS) languages adjunction to the Spec of CP is not available, the contrast between Bulgarian and the other Slavic languages concerning the possibility of multiple extraction from embedded clauses is derived.

## 3. Some flaws in Rudin's account: the concept of "Barrier" (Chomsky 1986).

According to Chomsky (1986), Subjacency is violated whenever two or more barriers are crossed<sup>2</sup>. One of such barriers can be "weak". The concept of "weak barrier" is parametrizable, and it is restricted not only to the Bounding Theory, but also to the most-deeply embedded clause. In English, for instance, the most-deeply embed-

(1) Rudin does not analyze Slovak but Slovak Comp should be assigned a structure like (8).

(2) The definition of "barrier" given by Chomsky (1986) is presented in (i):

- (i) "x" is a barrier for "y" iff (i) or (ii)
- (a) "x" immediately dominates "z", "z" a blocking category for "y".
- (b) "x" is a blocking category for "y", "x" not SF.

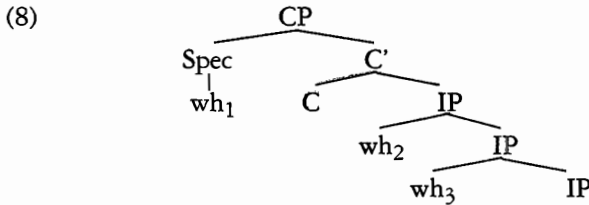
The definition of "blocking category", as well as the other relevant definitions are given in (ii)-(v):

- (ii) "x" is a blocking category for "y" iff "x" is not L-marked and "x" dominates "y".
- (iii) "x" L-marks "y" iff "x" is a lexical category that theta-governs "y".
- (iv) "x" theta-governs "y" iff "x" is a zero-level category that theta-marks "y", and "x" and "y" are sisters.
- (v) "x" is dominated by "y" only if it is dominated by every segment of "y". (Chomsky 1986).

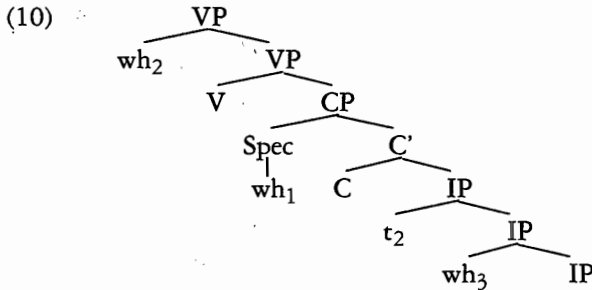
ded temporalized IP is a weak barrier, whereas in Italian or Spanish the weak barrier is the most deeply temporalized CP<sup>3</sup>.

Therefore, in order to account for certain cases of ill-formedness in languages where IP is a weak barrier through the Bounding Theory, it is necessary to preclude the possibility for *wh*-elements to adjoin to IP. Otherwise, not only the weak barrier, but also the barrier by inheritance would no longer exist<sup>4</sup>.

However, let us now consider again Rudin's structure for multiple questions in (-MFS) languages, repeated below:



In this configuration, it is clear that adjunction of *wh*<sub>2</sub>, for instance, to the immediately higher VP would not violate ECP: the trace left by *wh*<sub>2</sub> would be properly governed by its antecedent, since CP is not a barrier by inheritance (*wh*<sub>2</sub> is not dominated by all the segments of IP, so it is not dominated by IP) and, provided that V L-marks it, is not an inherent barrier either:



(3) This accounts for the contrast between (i) and (ii), as well as for the contrast between (i) and (iii) (already pointed out by Rizzi 1982):

- (i) \* Which car<sub>i</sub> did you wonder CP[to whom<sub>j</sub> IP[Tom gave t<sub>i</sub> t<sub>j</sub>]]
- (ii) Which car<sub>i</sub> did you wonder CP[how<sub>j</sub> IP[to fix t<sub>i</sub> t<sub>j</sub>]]?
- (iii) Il solo incarico che<sub>i</sub>; non sapevi CP [a chi<sub>j</sub> IP[avrebbero affidato t<sub>i</sub> t<sub>j</sub>]]...  
the only task that not knew-2sg to whom would have-3p1. assigned  
'The only task that you did not know to whom they could have assigned...'

In (i), the *wh*-element *which car* has crossed the most-deeply embedded CP, which is a barrier by inheritance; however, since the most-deeply embedded temporalized IP is also a barrier in English, two barriers have been crossed, thus violating Subjacency.

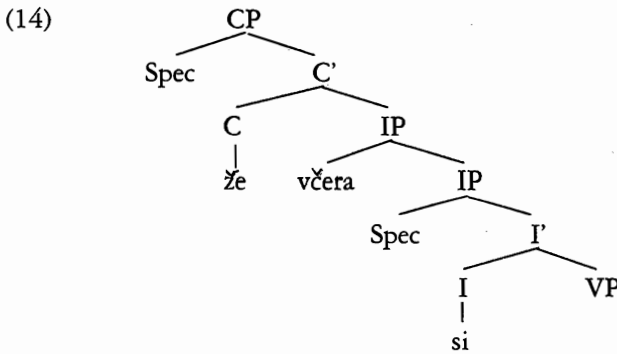
In (ii), the most-deeply embedded IP is not temporalized. Thus, since it is not a barrier, the *wh*-element has crossed just one barrier, namely, the most-deeply embedded CP (barrier by inheritance).

This is also the case in (iii), since in Italian it is the most-deeply temporalized CP, and not the IP, that counts as a barrier.

(4) The reason for this is that we are adopting May's (1985) assumption concerning adjunction structures. In a structure like (i), B consists of two "segments" and a category is dominated by B only if it is dominated by both of these segments:

- (i) B[ A B[...]]

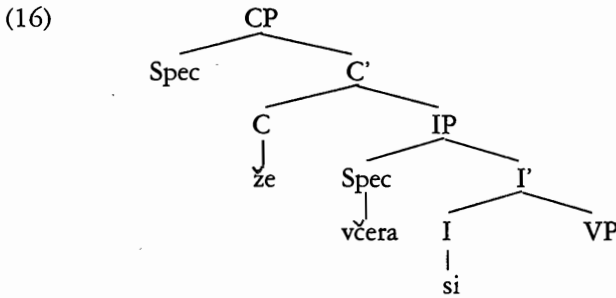




It is interesting to point out that the presence of a further constituent between *si* and *že* leads to ungrammaticality:

(15) \* *Myslím, že včera Peter si toto kúpil*

However, (14) should allow (15), since in (14) the position of Spec of IP remains empty, and it could be filled by the subject *Peter*. Since this is not possible, it will be necessary to assume that *včera*, though an adjunct, is filling the position of Spec of IP. Therefore, this position should be redefined as an A'-position so as to make it possible for the adverb to appear there (for the definition of the Spec of IP as a A'-position, see Pollock 1985, Santorini 1989 and Pesetsky 1989). Thus, the structure would be (16):



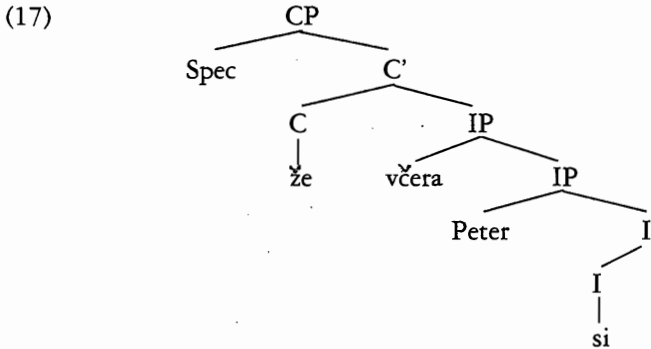
Unfortunately, this hypothesis is not compatible with recent proposals about Scrambling (cfr. Grewendorf and Sternefeld 1989), according to which Scrambling is adjunction to IP or VP (some authors even reject adjunction to VP; see Fanselow 1989). Free word order in Slovak can certainly be attributed to a Scrambling process<sup>5</sup>. However, proposals such as (14) or (16) (no matter whether or not *včera* occupies the subject position) presuppose that Scrambling cannot consist of adjunction to IP. Indeed, if this were the case, once adjunction to IP is done, an unrestricted

(5) The following examples, all grammatical, will give the reader an idea of the freedom in the Slovak word order:

(i) a. *Otec Jankovi vždy vysvetľuje všetko.*  
 father to Janko always explains everything  
 "The father always explains everything to Janko".

b. *Otec vždy Jankovi všetko vysvetľuje.*  
 c. *Vždy Jankovi otec vysvetľuje všetko.*  
 d. *Jankovi všetko otec vždy vysvetľuje.*

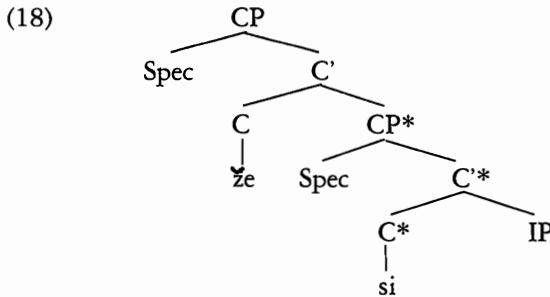
number of constituents could appear between *že* and *si* in (13). This would incorrectly predict the grammaticality of a structure such as (17), which corresponds to (15):



Let us now suppose that *si*, in (13), is not in Infl, but in C. It is not the same C where *že* is placed, since, as has been already said, this would not allow us to understand which is the position filled by *včera*. It is a case of Comp recursivity, as we will see in the next section.

#### 4.2. Clitics in C: Comp recursivity in Slovak.

We propose that (18) is the correct structure for (13):

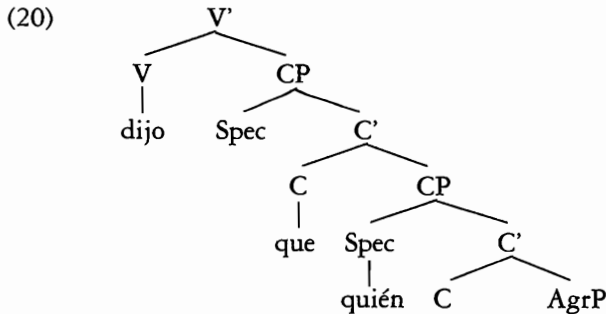


Notice that is a case of Comp recursivity. The existence of Comp recursivity is valid within the framework of Chomsky (1986): a complementizer has the ability of subcategorizing for a further CP as its complement.

A first piece of evidence in favor of Comp recursivity is given by Plann (1982). Plann suggests that Spanish indirect questions where a complementizer *que* 'that' precedes the *wh*-element, as in (19), have to be represented by means of CP recursivity, as in (20)<sup>6</sup>:

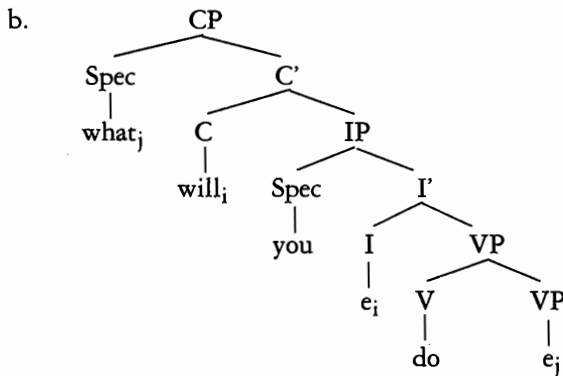
(6) For similar phenomena concerning Hungarian, see Marácz (1989), where it is also argued that CP recursivity is at work in such cases.

- (19) Juan dijo que quién viene.  
 Juan said that who comes



A second piece of evidence in favor of Comp recursivity is due to Rizzi and Roberts (1989). According to Rizzi and Roberts, a sentence such as (21a) (a case of English subject-verb inversion) can be understood as a case of verb movement by the auxiliary *will* from its base position into C, as shown in (21b):

- (21) a. what will you do?

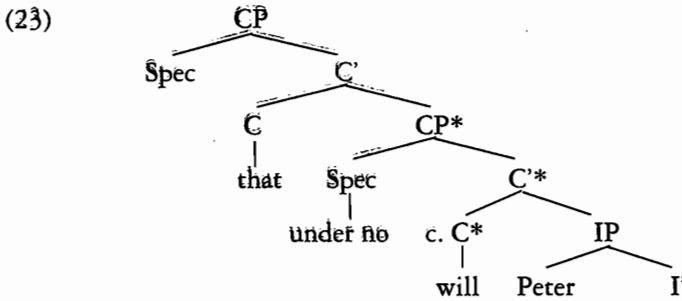


The movement of *will* from Infl to C is forced by a sort of "Wh-Criterion" (cfr. May 1985): the presence of a wh-element in the Spec of CP requires that the head C also share wh-features. Rizzi (1990b) suggests that Tense morphemes bear such features, so that movement of the node Infl to C is obligatory in order to provide C with the wh-features.

Afterwards, Rizzi and Roberts consider a sentence such as (22), where the English subject-verb inversion phenomenon takes place in an embedded sentence whose C position is already filled by a complementizer *that*. So, they suggest that, in these cases, *that* is subcategorizing for a further CP (CP\* in (23)), thus triggering the appearance of a new head C (C\* in (23)) where *will* can land:

- (22) John thinks that under no circumstances will Peter leave.

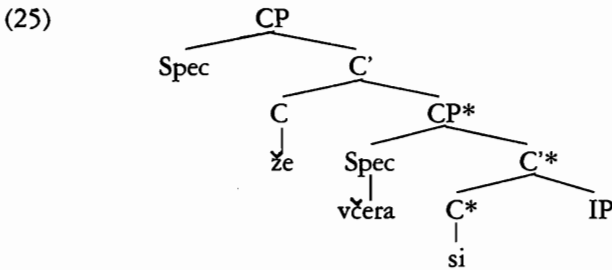




Rizzi and Roberts give a piece of evidence in favor of this configuration: the presence of *under no circumstances* in the Spec of CP\* gives rise to weak island effects, as shown in (24), where the internal argument of the embedded clause (*what*) has been moved to the matrix clause:

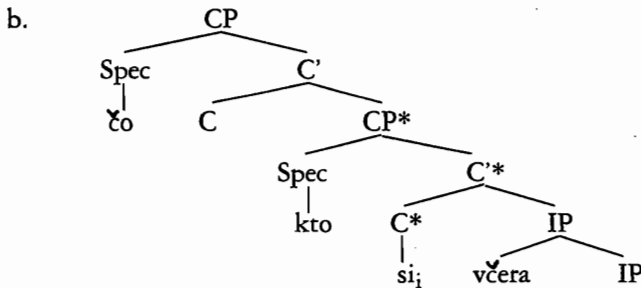
(24) \* What did he say that under no circumstances would he do?

Therefore, if the Comp recursivity strategy is available, we propose that, in (13), *včera* occupies the Spec of a CP subcategorized by a higher C (CP\*), and *si* occupies C\*, as shown in (25):

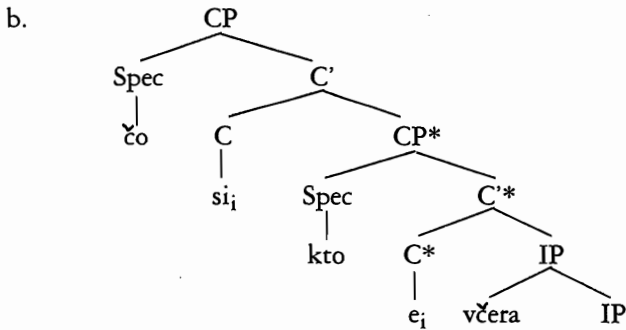


It is not possible that *si* raises to C, since C is already filled by the complementizer *že*. However, if we suppose that Comp recursivity can also be possible in a root clause, *si* will have to raise further into the highest C for some reason. For instance, the grammatical sentence which corresponds to (26a) (whose structure is (26b)) is (27a), whose structure is (27b):

(26) a. \* čo kto si včera kúpil.



(27) a. *čo si kto včera kúpil?*



Let us suppose that the obligatoriness for the clitic to raise to the highest C is the same as the obligatoriness for the auxiliary to reach C in English subject-verb inversion sentences: provided that clitics are generated under Infl (a matter which will be the main topic of the next section), further clitic raising is actually Infl raising. Infl raising is done in order to provide all the successive CP Spec's with a corresponding C head endowed with the wh-features carried by the Tense morphemes under Infl, thus making possible the fulfillment of the Wh-Criterion. In (27b), the clitic (that is, Infl) has reached C\* because of the Wh-Criterion, and later has reached C for the same reason. By contrast, the ungrammaticality of (26a) is derived from the fact that C does not bear wh-features, since no movement of Infl to the C position has taken place.

This forces clitics in a matrix clause always to appear immediately to the right of the first wh-constituent, which is the salient difference between (26a) and (27a). However, thus far the reason is not at all clear why clitics are related to the wh-features which satisfy the Wh-Criterion, that is, why they have been related to Infl. The next section will be devoted to this topic.

### 4.3. Pronominal Clitics and the node Tense.

#### 4.3.1. Some further data about Slovak.

As said above, in Slovak, the constituent formed by the auxiliary verb and the unstressed pronouns, that is, the clitic cluster, has to occupy the "second position" in the sentence:

- (28) *Videl som ho včera.*  
 see-past. part. aux.perf.1sg. him yesterday  
 'I saw him yesterday.'

In (28), the past participle *videl* is in the first position, but any other constituent can occupy that place:

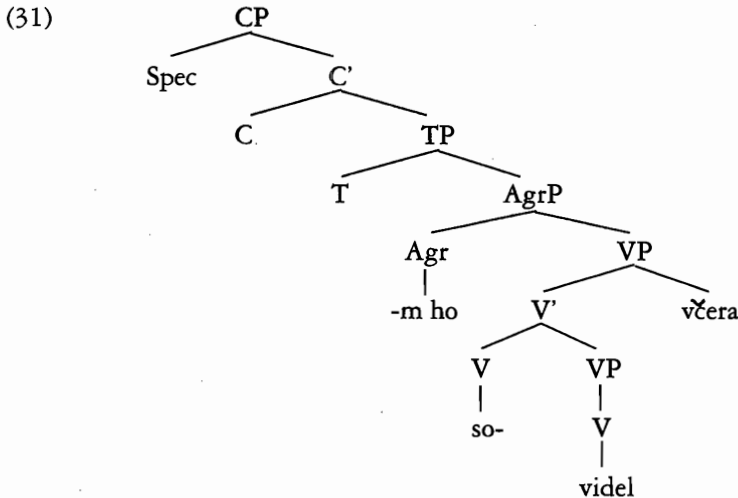
- (29) a. Ja som ho včera videl.                      b. Včera som ho videl.  
           I  
           'I saw him yesterday.'

It is not possible to put both *ja* 'I' and *včera* 'yesterday' before the clitic cluster, since this would violate the second position requirement:

- (30) \* Ja včera som ho videl.

We will account for these facts by considering the clitics to always be related to the position C, either by adjunction or substitution. Furthermore, from now on we will adopt Pollock's (1989) proposal about Infl, according to which Infl no longer exists as a unique projection, and is split in two different functional categories: an Agreement Phrase (AgrP) and a Tense Phrase (TP).

As far as (28) is concerned, let us propose (31) as the DS<sup>7</sup>:



At SS, V raises to the auxiliary *so-*, the auxiliary raises to Agr, Agr to T, and T to C, leading to the order in (28)<sup>8</sup>.

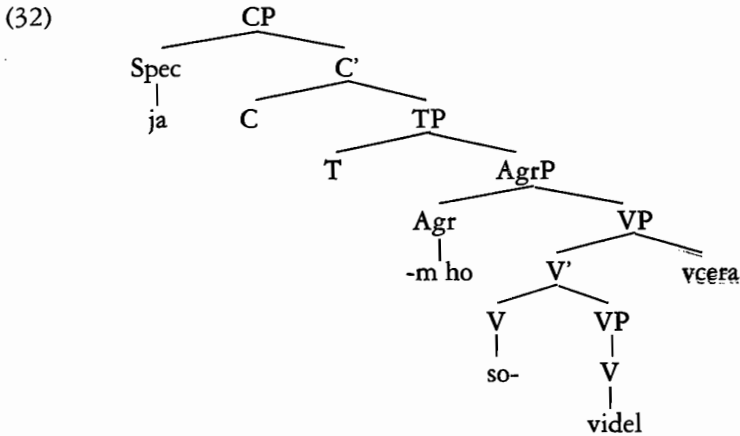
As far as (29) is concerned, we will consider (32) to be the correct representation (we refer only to (29a)):

(7) It will not be relevant for this article whether or not the temporal adverb *včera* really occupies the Spec of the VP headed by the auxiliary *som*. The position assigned to this adverb in (31) must not be interpreted as a proposal concerning this question, but rather as an arbitrary choice. At least on intuitive grounds, temporal adverbs would be more likely to actually appear as adjuncts to TP.

As for the structure of V', we have simply assumed Chomsky's (1986) hypothesis about the relation between auxiliaries and theta-role assigning verbs.

(8) That *videl* takes part of C in (28) accounts for the impossibility of inserting any constituent between V *videl* and *som* whenever *videl* is the leftmost word in the sentence.

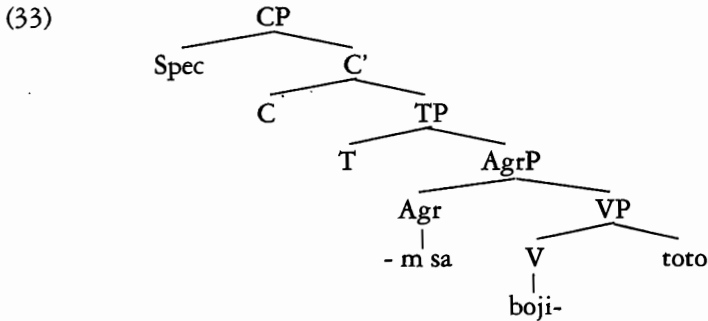
(i) \* Videl včera som ho.



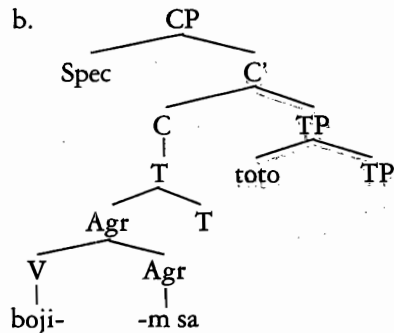
In (32), the auxiliary *so-* raises to Agr, Agr raises to T and T to C, giving the order in (29a) (in section 4.5, we will explain why V raising is impossible when the Spec of CP is already filled).

#### 4.3.2. T-to-C movement: clitics and the Tense operator.

In (31) and (32), the clitic *ho* appears under Agr, behind the person-number morpheme. This predicts that V-to-Agr-to-T-to-C movement gives rise to a final order of the type V-Agr-clitics-T. Indeed, taking (33) as the starting point, V-to-Agr-to-T-to-C movement gives rise to (34a), whose final representation is (34b):

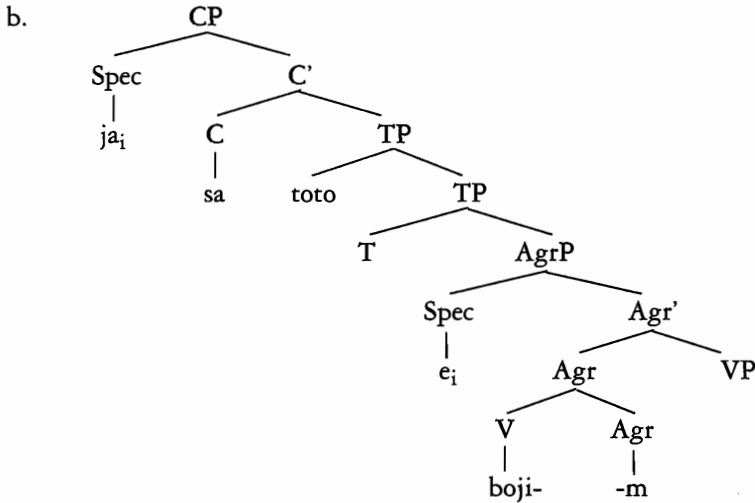


(34) a. Bojim sa toto.  
fear-1sg. reflex. this  
'I fear this.'



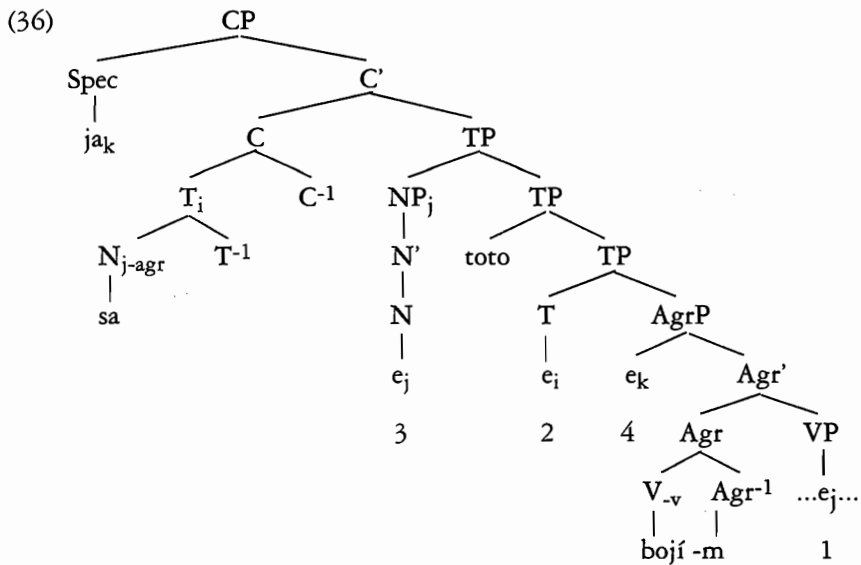
However, (35a) (whose structure is (35b) raises a problem: how can the clitic *sa* reach C alone, if every  $X^0$  node blocks the extraction of any element which occurs under it (Kayne 1989)<sup>9</sup>:

- (35) a. Ja sa toto bojím.  
'I fear this.'



This means that *sa* never landed on either V or Agr. Let us suppose that these pronominal clitics reach T directly, without passing through V and Agr (a process which clearly contrasts with the step-by-step movement proposed by Kayne (1989). The empty category left behind would be identified without problems, since the V-to-Agr movement allows Agr to L-mark VP and AgrP can no longer be a barrier by inheritance. Nevertheless, a problem with Relativized Minimality (Rizzi 1990a) arises if this account is accepted. Clitics are heads, since they land on T. However, other heads (V, Agr) intervene between such clitics and their original traces. This would not allow the clitics to identify their traces because of Relativized Minimality. This problem can be circumvented in the following way. A clitic is the head of a maximal projection. The movement of such a head takes place not from the DS position (where the maximal projection is generated), but from the position that the maximal projection reaches by Scrambling (as said, Scrambling is adjunction to IP, that is to say, to TP in Pollock's 1989 syntactic tree). Therefore, at a certain step of the derivation we have the situation in (36):

(9) In (35), *bojim* 'fear' is not in C, as the occurrence of *toto* 'this' reveals.

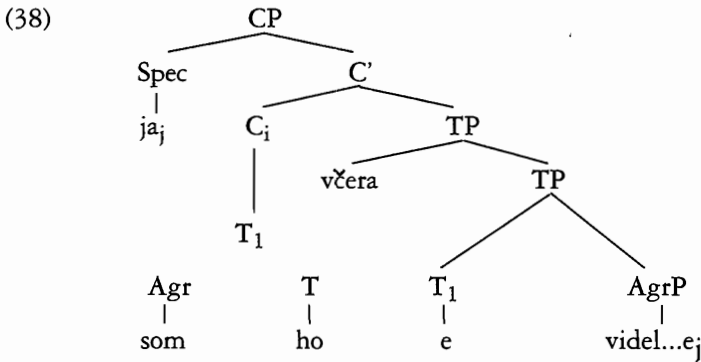


- 1: trace resulting from Scrambling ( $NP_j$  adjoins to TP).
- 2: trace resulting from T-to-C movement (obligatory at SS).
- 3: trace resulting from N-to-T incorporation (substitution).
- 4: trace resulting from the movement of  $ja_k$  to the Spec of CP.

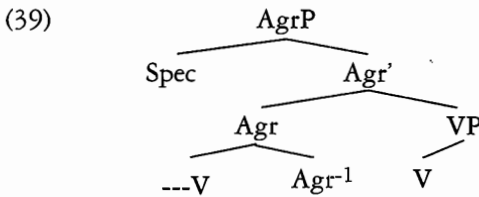
We have seen in the first section that, in Slovak, all the wh-operators are moved at SS. In those languages where a wh-operators is not moved at SS, the movement is performed at LF (cfr. Huang 1982 for Chinese, where no wh-movement is possible at SS). The landing site for wh-operators at LF is the Spec of CP. Let us now suppose that the Tense morpheme is also an operator, and that its landing site at LF is C. This is not an original idea. Raposo (1986), for instance, proposes that a [+ TENSE] operator occurs in Comp which binds the tense morphemes at a distance in Romance languages. More recently, Rizzi (1990b) has argued that the node T can bear wh-features, and that the movement of T to C is crucial for the fulfillment of the Wh-Criterion at SS in English. As far as Slovak is concerned, let us suppose that, as in the case of Slovak wh-operators, Slovak T has to perform at SS what operators in other languages perform at LF, namely, it has to raise to C. Briefly, operators in situ do not exist in Slovak. This is the reason for the contrast between (37a) and (37b):

- (37) a. Kto čo robil?  
 who what did  
 'Who did what?'  
 b. \* Kto robil čo?

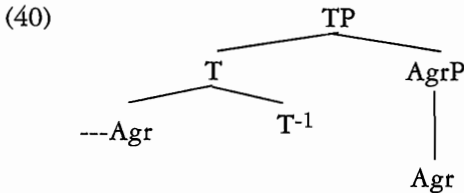
V-to Agr-to-T-to-C movement takes place, giving rise to sentences such as (34a). In (35a), V did not raise to T (as said above, in section 4.5. we will explain why V raising is impossible when the Spec of CP is already filled). In these cases, however, auxiliaries can appear under T. This is illustrated in (29a), whose structure is (38):



Why does T attract the clitics? Rizzi and Roberts (1989) suggest that head-to-head movement is actually substitution, not adjunction. A negative projection (*à la* Selkirk 1982) of a head  $X^0$  (that is,  $X^{-1}$ ) subcategorizes for a slot which a second head Y (the head of the maximal projection subcategorized by the former head) can fill. For example, in (39), Agr subcategorizes for a maximal projection VP, so that  $\text{Agr}^{-1}$  can project a position which the head of VP, V, can fill:



In (35), T subcategorizes for AgrP, so that the tense morpheme also subcategorizes for a place which Agr will be able to fill. This is illustrated in (40):

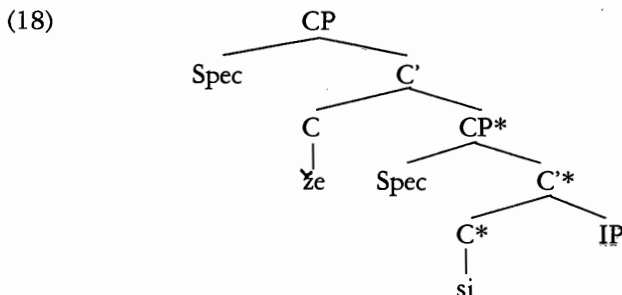


If pronominal clitics are defined as agreement morphemes<sup>10</sup>, it is reasonable to think that their landing site should also be under T. Therefore, the invariable second position of Slovak clitics is due to the movement of T to C at SS.

(10) An anonymous reviewer of *ASJU* has pointed out to me that the characterization of these pronominal clitics as agreement morphemes seems to imply that an Object Agreement functional projection is also playing a role in the structures proposed so far. Indeed, I do not discard the possibility that such a functional category is present in the Slovak syntactic structures. However, the claim that pronominal clitics are agreement morphemes is rather based on Torrego's (Spring lectures, Madrid 1990) proposal that pronominal clitics are generated in the head of an agreement projection located *inside* an NP (which actually is a DP, according to Abney 1987). Later, this head has to undergo incorporation into another head. Our hypothesis is that, for Slovak, the "agr" slot under T is the landing site for that incorporation process.

#### 4.4 Some further advantages of Comp recursivity.

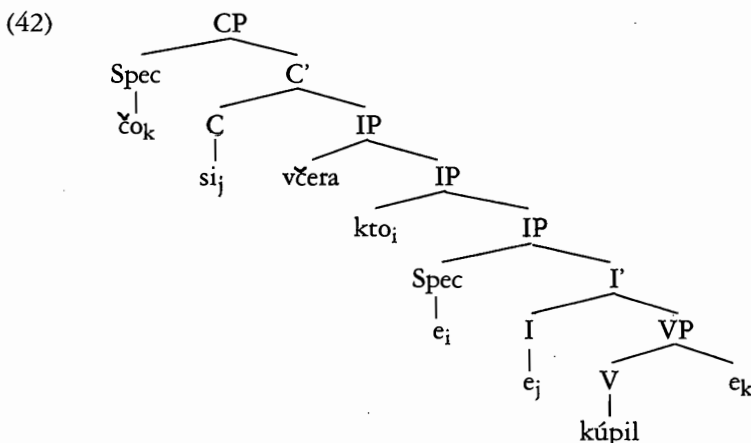
Notice that, by means of a proposal based on Comp recursivity, we can account for the fact that, in embedded clauses, just one constituent can appear between *že* and *si*: as is clear in (18) (repeated below), under Comp recursivity just one slot (the Spec of CP\*) is available for a constituent to raise to:



On the other hand, (27b) is compatible with the widespread conception of Scrambling as a process of adjunction to IP, since, in this case, all the adjoined constituents obligatorily reach a position to the right of the clitic. Moreover, they have also to reach a position to the right of the wh-operator. Indeed, examples such as (41) are ungrammatical:

- (41) \**čo si včera kto kúpil?*  
 what to himself yesterday who buy-past-3sg.  
 'Who bought what yesterday?'

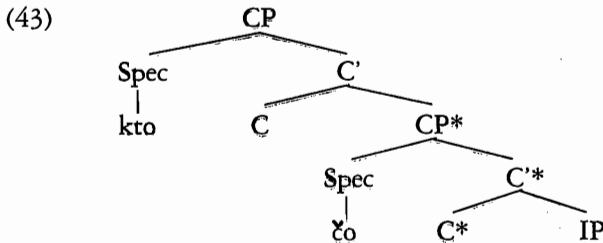
Notice that Rudin's (1988) hypothesis, namely (8), would predict the grammaticality of (41) since, according to this, wh-operators as well as scrambled constituents would be adjuncts to IP, as (42) show:



Third, by means of the Bounding Theory, (27b) can be useful to give a correct answer for the impossibility of simultaneously extracting several wh-elements from



embedded clauses in (-MFS) languages. If we consider that, in such languages (like in Spanish or Italian, pro-drop languages as well), the "weak" barrier is CP, rather than IP, a structure such as (43) will exhibit two barriers (CP and CP\*) for the extraction of a wh-element which has already reached the Spec of CP\*:



Fourth, our hypothesis based on Comp recursivity fits better the notion "parameter". Rudin's article focusses on the possibilities of parametrization. In this case, the parametrized principle concerns adjunction, and it is called the Condition on the Spec-CP Adjunction (CSA), presented in (44):

- (44) Condition on SpecCP Adjunction  
 \*SpecCP[  $\alpha$  SpecCP ]  
 (nothing may be adjoined to SpecCP)

Recently, it has been proposed (Manzini and Wexler 1987) that reducing the relevant parameters to lexical particularities would remarkably simplify hypotheses about language acquisition. On the other hand, it has also been pointed out that languages can differ in the fact that a functional category can select one or another functional/lexical category<sup>11</sup>. (44) cannot be related to this general trend, since no property of lexical or functional heads is at work. By contrast, in our proposal, it is a particularity of the morphemes in C (in this case, the presence or absence or the property of subcategorizing for a further CP) that discriminates (-MFS) languages from (-MFS) languages. In Section 4.5 we treat in more detail the properties of such morphemes, and in Section 4.6 we will claim that the basic contrast between Slavic languages has to do not with the availability of the Spec of CP as an adjunction site, but rather with the availability of Comp recursivity.

#### 4.5. Properties of C in Slovak.

##### 4.5.1. Weak Agreement in C.

In Sections 4.3.1. (example (28)) and 4.3.2. (example (34)) we talked about V-to-C movement. In those examples, the Spec of CP had to remain empty:

- (45) a. \*Včera videl som ho.      b. \*Ja bojím sa toto.

In this section we will try to give an account for this fact, since pursuing this topic will allow us to find out a relevant property of C in Slovak. The account is based on the ECP.

(11) For instance, Laka (1989) claims that the functional category Negation Phrase can be base-generated under T, as in French and English, or above T, as in Basque.

We assume Rizzi's (1990a) conjunctive formulation of the ECP:

- (46) A non-pronominal empty category must be:  
 a. head governed (Formal Licensing)  
 b. antecedent-governed or Theta governed  
 (Identification)

Furthermore, Rizzi proposes the following expansion for the node C in English wherever CP is [+finite]:

- (47) C(T) → (that) / (Agr)

Both expansions are in complementary distribution, and are optional. (47) allows us to account for English "that-trace" effects in the following way: when expansion is *that*, a trace in subject position cannot be head-governors (48a); when Agr is in C, head-government is possible (48b); when no expansion is chosen, C remains empty and, although an auxiliary may occupy this position, this movement cannot provide C with the property of being head-governor (48c):

- (48) a. \*Who<sub>i</sub> do you think [<sub>i</sub> that[e<sub>i</sub> left]]?  
 b. Who<sub>i</sub> do you think [e'<sub>i</sub> Agr[e<sub>i</sub> left]]?  
 c. \*Who<sub>i</sub> did [e<sub>i</sub> leave]?

Rizzi says that a constituent has to occur in the Spec of CP in order to license Agr in C: in (48b), for instance, a trace (e'<sub>i</sub>) occurs in the Spec of CP.

(47) allows to account for (45) if it is assumed that Slovak Agr in C is "weak" in Pollock's (1989) sense, that is, the incorporation of V to such an Agr would not allow V to transfer the ability of assigning theta-roles to its trace, thus giving rise to a Theta-Criterion violation<sup>12</sup>.

Let us now suppose that, in Slovak, it is not necessary for Agr to occur in C when Spec of CP is filled. That is, let us strictly follow (47) in accepting that the null expansion is available in this language too. In this case, the ungrammaticality of (45) would not yet have been accounted for.

Therefore, we will assume that, in Slovak, the null expansion of C is never possible when the Spec of CP is filled by some constituent. The difference with respect to English can be based on a more general phenomenon which has to do with the contrast between languages which can exhibit no-operator constituents in the Spec of CP, and languages which can not. Rizzi (1990b) accounts for this contrast by means of a features system to define functional categories. The two relevant features are [ $\pm$ C] and [ $\pm$ I]. [+C] means "propositional", and [+I] means "predicational". The system gives rise to the following categories:

- [-C, -I]: a functional category neither predicational nor propositional (perhaps DP);  
 [+C, -I]: a category which designates a proposition (CP in non-V-2 languages and in non-V-2 clauses of V-2 languages);  
 [-C, -I]: a category meaning predication (IP);

(12) This parameter concerning agreement morphemes explains why no V movement can take place in English, as well as the possibility of V movement in French finite clauses (cfr. Pollock 1989).

[+C, -I]; a category meaning predication and proposition. It corresponds to the V-2 clauses in V-2 languages.

The existence of a CP with the values [+C, -I] in V-2 languages (as in German) accounts for the fact that both operators and non-operators can appear in the Spec of that CP. Operators are licensed by a C with the value [+C], and non operators are licensed by the value [+I]. By contrast, in English, where C is [+C, -I], the Spec of CP can only be filled by operators. In Slovak, since operators as well as non-operators can appear in the Spec of CP, it will also be the case that the features in C are [+C, -I].

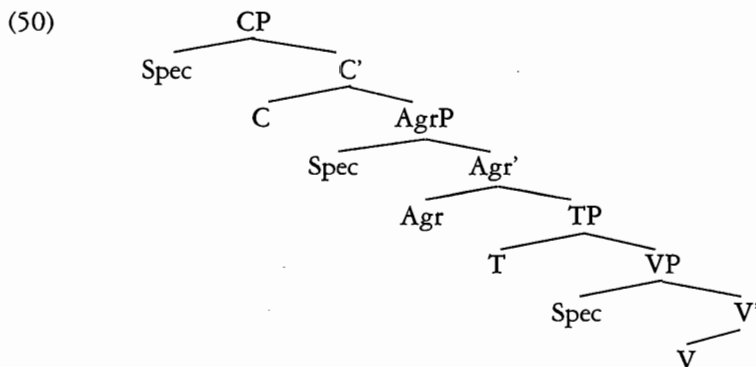
Now, suppose that the presence of the value [+I] for C correlates with the unavailability of the null expansion in (47). So, the presence of a constituent in the Spec of CP, whether operator or non-operator, will necessarily co-occur with Agr morphemes in C. These morphemes are always able to head-govern, whence the grammaticality of a German sentence such as (49a), whose structure is (49b):

- (49) a. Maria kam.  
'Maria came.'
- b. CP[<sub>Spec</sub>[Maria<sub>i</sub>] C[<sub>kan</sub>] IP[<sub>t<sub>i</sub> t<sub>j</sub> ]]</sub>

This is a minimal departure from Rizzi's account for the grammaticality of (49a). Rizzi proposes that (49a) is wellformed because the only head that can head-govern the subject variable is the Infl which raises, and in (49) the feature [+I] of the raising Infl matches the feature [+I] which already existed in C, unlike what happened in English, where there was no value [+I] in C at all. However, according to our account, it is the Agr in C, rather than the raising Infl, that head governs the subject variable. This allows us to maintain the idea that, in languages where the CP shares the features [+C, -I], C will host Agr whenever the Spec of CP is filled by some constituent.

In Slovak such an Agr in C is weak. This means that the lexical V will never be able to raise to C unless C lacks Agr, something that only happens when the Spec of CP is empty. In German root clauses, the lexical V can always appear under C (as in (49a)). This suggests that, in this language, such Agr morphemes in C are strong.

Thus far, we have not yet accounted for the reason why the German V (whether lexical or not) has to appear in C in root clauses, unlike Slovak V. Let us suppose that this has to do with the different hierarchy or functional categories in both languages. In German, the hierarchy is as follows:

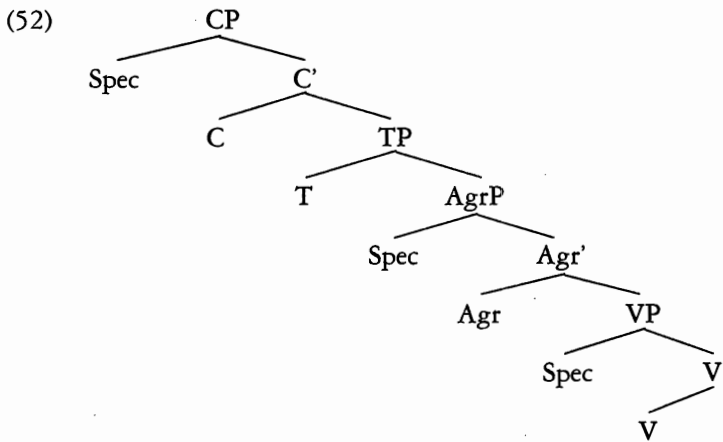


Since V has to raise to form a cluster with Agr, it also has to cross T in order to reach Agr. The T morphemes, then, remain inside the cluster V-T-Agr. However, for some reason, the T morphemes cannot stay under the node Agr at SS. Indeed, Rizzi (1990b) adapts a principle proposed by Laka (1989) (the Tense C-Command Condition) in the following way:

- (51) The tense specification must c-command all the other [+I] categories in the same clause.

This principle forces T to move to C (which has a [+I] value). However, since T takes part of the cluster V-T-Agr, it is the whole cluster that has to move to C.

In Slovak the situation is quite different. The hierarchy of functional categories is (52):



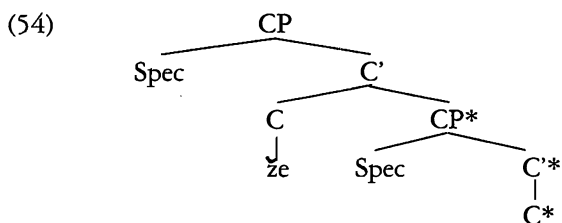
In this case, V raises to Agr, forming the cluster V-Agr. But T remains free, and it can raise alone to C leaving behind the cluster V-Agr. Let us also point out that the movement of T to C in Slovak, which we attributed in Section 4.3.2. to the impossibility of having operators in situ in such a language, can now be rephrased in Laka/Rizzi's terms: movement of T to C is due to (51).

#### 4.5.2. CP Recursivity and expletive Agr in C.

Let us consider the following sentences, which exhibit an embedded clause:

- (53) a. Povedal som, že som kávu pil včera.  
 say-pres.part. aux.perf.1sg. that aux.perf.1sg. coffee drink-part.pres. yesterday  
 'I said that I have coffee drunk'.  
 b. \* Povedal som, ze pil som kávu včera.

We are proposing that *že* 'that' is able to subcategorize for a further CP, as illustrated in (54):



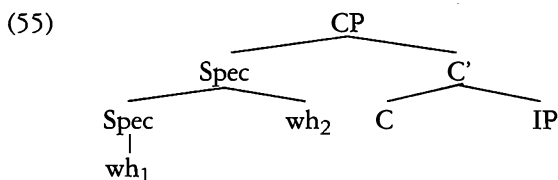
As in (28), *som* fills C in (53). However, we do not yet understand why, in case the Spec of CP\* is not filled by any constituent, V cannot raise to C\* in non-root contexts. This should be possible, since we have said that the absence of a constituent in the Spec of CP\* would allow C\* to remain empty.

We propose that, in these cases as well, the head of CP\* has certain features carried by certain morphemes. The idea is that, just as in the case of embedded questions, where a CP subcategorized by a verb like *wonder*, for example, has to bear [+wh] features, selection by C is also coupled with the appearance of certain morphemes which fill the head of the selected constituent. However, something has to identify and formally license such morphemes. We here suggest that they are expletive and that its identification is made possible by *ze* itself. We further suggest that they are weak too, so that V-to-C raising would lead to a Theta-Criterion violation, hence the ungrammaticality of (53b).

#### 4.6. Multiple Questions in Bulgarian.

In this section, we will propose that the parameter responsible for the splitting of Slavic languages in two groups has nothing to do with (44), but rather with the (un)availability of Comp recursivity.

Let us suppose, indeed, that Bulgarian lacks Comp recursivity. This does not lead us to accept Rudin's proposal, which says that multiple preposing in Bulgarian multiple questions is due to a successive adjunction (to the right) of wh-elements to the Spec of CP, as shown in (55):



Several problems arise with this approach.

First, what Rudin calls adjunction to the Spec of CP is actually adjunction to a wh-element, that is, in many cases it is adjunction to a NP. NP is an argument category (cfr. Sportiche 1989) which, as such, does not accept that any other constituent adjoins to it, according to Chomsky (1986)<sup>13</sup>.

(13) A reviewer of *ASJU* has pointed out that Chomsky's (1986) restriction on the kind of categories which a constituent can be adjoined to is based on data where those categories occupy an argument position, which is not the case in (55). As a result, such a fact could undermine this first objection. However, this is not the case if we accept Sportiche's (1989) approach to this question, according to which the fact of whether or not the argument category is occupying an argument position is not relevant.

Second, Bulgarian exhibits a strict linear order between the different wh-elements: subject/direct object/indirect object, as illustrated in (56):

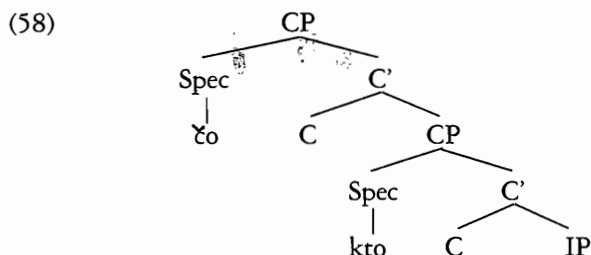
- |                                     |                          |
|-------------------------------------|--------------------------|
| (56) a. koj kogo na kogo e pokazal? | c. * na kogo koj kogo... |
| who whom to whom has pointed out    | d. * na kogo kogo koj... |
| 'Who pointed out whom to whom?'     | e. * kogo koj na kogo... |
| b. * koj na kogo kogo...            | f. * kogo na kogo koj... |

Rudin attributes the relative order subject/internal arguments to the Superiority Condition. This view has to face two problems: on the one hand, it conflicts with the fact that Superiority effects do not exist in other Slavic languages (as in Slovak for instance); on the other hand, Rudin cannot offer an account for the strict linear order between direct and indirect objects. The following paragraphs will be devoted to these two problems.

The relative order subject/internal arguments is accounted for by Rudin in the following way. According to her, the wh-subject should occupy the Spec of CP because it must transmit its index to C through Spec-head agreement. The reason is that, under Aoun, Hornstein, Lightfoot and Weinber's (1987) framework (AHLW) (assumed by Rudin), it is C that has to bind the variable left by the wh-element. However, in principle, nothing should preclude, in Bulgarian, the wh-subject from adjoining to IP before reaching the Spec of CP, since Rudin has also assumed the availability of such an adjunction in (-MFS) languages. If this is the case, the variable adjoined to IP, and this variable, lacking a domain, would simply need to be bound by a quantifier. In AHLW's framework the crucial fact is that it is not the quantifier that is binding the variable-subject, but rather C, which the quantifier has transmitted the index to. Otherwise, the variable could be bound by the quantifier even in those cases where the quantifier is adjoined to the Spec of CP (thus not transmitting its index to C), since in this configuration the quantifier would c-command the variable (the first branching node which dominates the quantifier is CP, and not the NP projection dominating the adjoined element, since it is just one of the two segments which form the whole NP projection). Evidence in favor of this possibility is the fact that, in a configuration where the Superiority effect has not taken place, the requirement still stands that a variable without a domain be related to a quantifier, which can only be possible through c-command. Therefore, if this is the case, the account for the Superiority effects in Bulgarian would need to claim either that the traces adjoined to IP cannot bind, or that such traces do not exist in Bulgarian. Since the latter alternative would amount to establishing an exception within the Bounding Theory in Rudin's terms (recall that she accepts adjunction to IP for the [-MFS] languages, hence also the existence of traces in those positions), the former will have to be chosen. Moreover, this alternative agrees with AHLW's idea that a variable can only be bound within its domain by a C coindexed with the wh-element which is the antecedent of the variable. This approach would allow us to preserve Rudin's proposal for Bulgarian, but it would not account for why a sentence such as (57) is possible in Slovak:

- (57) Čo kto robil?  
 what who did  
 'Who did what?'

According to Rudin (recall (8)), in (57) *čo* appears in the Spec of CP and *kto* is adjoined to IP. Indeed, the subject variable is not bound by a C coindexed with the corresponding wh-element. By contrast, a structure such as (58) does provide the variable with such a C, thus enabling us to account for the lack of Superiority effects in (-MFS) languages:



Summarizing, the proposal by Rudin in (55) for the Superiority effects in Bulgarian leaves (58) as the only alternative to account for the lack of Superiority effects in (-MFS) languages. This, of course, is not a desirable result for Rudin<sup>14</sup>.

The second problem arising under the ALHW-based account for (56) are the above mentioned restrictions between direct and indirect object wh-elements, as well as the fact that Bulgarian multiple questions do not exhibit wh-elements non-subcategorized by the verb (the equivalent to *why* and *how*).

Let us assume the following points:

- a) Bulgarian, like Slovak, also exhibits Scrambling, this process consisting of an adjunction to IP (examples like the ones in note 5 are common in Bulgarian);
- b) wh-elements cannot undergo Scrambling<sup>15</sup>;
- c) Bulgarian verbs raise to Infl from their base position, leaving behind the constituent VP (so, Bulgarian is like most Slavic languages in this respect).

As a result of these three assumptions, only the interrogative constituents would remain within the VP. Let us now suppose that the presence of several wh-elements at the beginning of the Bulgarian sentence is not due either to (44) or to the existence of Comp recursivity, but rather to the movement of a constituent  $V^{\max}$  (cf.

(14) Of course, the counter-argument just presented can work just in case an analysis of [-MFS] languages based on CP recursivity is rejected. In itself, as a reviewer of *ASJU* observed, Rudin's proposal for Bulgarian seems to precisely correlate Superiority effects and properties of multiple questions in this language, but it is interesting to point out that her hypothesis for [+MFS] languages in the way just presented.

(15) Grewendorf and Sternefeld (1989) say that this generalization is illustrated by German multiple questions like the following:

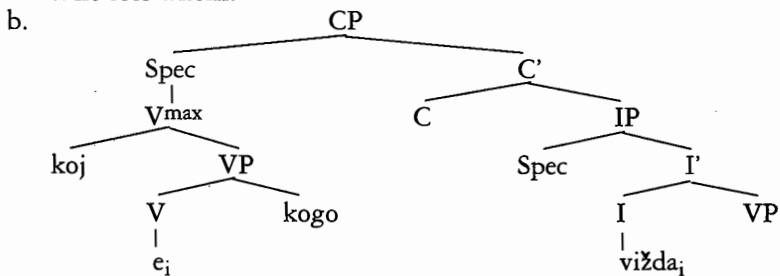
- (i) a. *Wem hat der Student welche Frage beantwortet?*  
 who(m) has the student which question answered  
 'Who did the student answer which question?'  
 b. \*? *Wem hat welche Frage der Student beantwortet?*
- (ii) a. *Warum hat jeder welches Buch gekauft?*  
 why has everyone which book bought  
 'Why has everyone bought which book?'  
 b. \*? *Warum hat welches Buch jeder gekauft?*

Sportiche 1988) to the Spec of CP<sup>16</sup>. A similar proposal has already been done by den Besten and Webelhuth (1987) for German<sup>17</sup>, although, in this case, the maximal projection of V cannot dominate anything but the V itself, which does not raise to Infl either when it cooccurs with an auxiliary, or in non-root contexts, as in (59):

- (59) Gegessen habe ich viele Apfel.  
 eaten have I many apple  
 'I have eaten many apples'.

Thus, the structure of a sentence such as (60a) will be (60b):

- (60) a. koj koj vižda  
 who whom sees  
 'Who sees whom?'



The movement of  $V^{\max}$  to the Spec of CP is possible because the trace left is properly governed by Infl (for a similar case in English, cfr. Roberts 1988). The trace  $e$  of (60b) is antecedent-governed by *vižda* once Reconstruction has been done at LF. On the other hand, we have accepted Koopman and Sportiche's (1985) hypothesis about subject position: *koj*, by moving to Spec of CP, gives rise to pied-piping of the whole  $V^{\max}$  where it has been generated.

Summarizing, the strict linear word order between wh-elements in Bulgarian mirrors their base position within  $V^{\max}$ .<sup>18</sup>

If we are on the right track with respect to Bulgarian, the clitic which raises into C will stop to the right of  $V^{\max}$ , that is, to the right of the whole set of wh-elements. This position of the clitics with respect to wh-elements is one of the salient differences between [-MFS] and [+MFS] languages. (61a,b) illustrate this fact:

- (61) a. koj kakvo ti e kazal?  
 who what you aux.perf.3sg. told  
 'Who told you what?'  
 b. \* koj ti e kakvo kazal?

(16) The category  $V^{\max}$ , up to now ignored in this paper, but compatible with all the ideas we have developed so far, will be relevant for the proposals to be presented in the next paragraphs. The existence of such a category has been defended by Sportiche (1988), and it can be conceived as a kind of small clause. The external argument of the verb is generated in the Spec of  $V^{\max}$ , what allows Sportiche to account for the behavior of floating quantifiers in French.

(17) Let us point out that CP piping, a similar process to  $V^{\max}$  pied-piping, has also been suggested for Basque questions in Ortiz de Urbina (1988).

(18) The fact that Basque, a language which has Scrambling, also displays a strict linear order for wh-elements in multiple questions, suggests that something like  $V^{\max}$  or CP (see footnote 17) pied-piping has to be at work in this language as well.

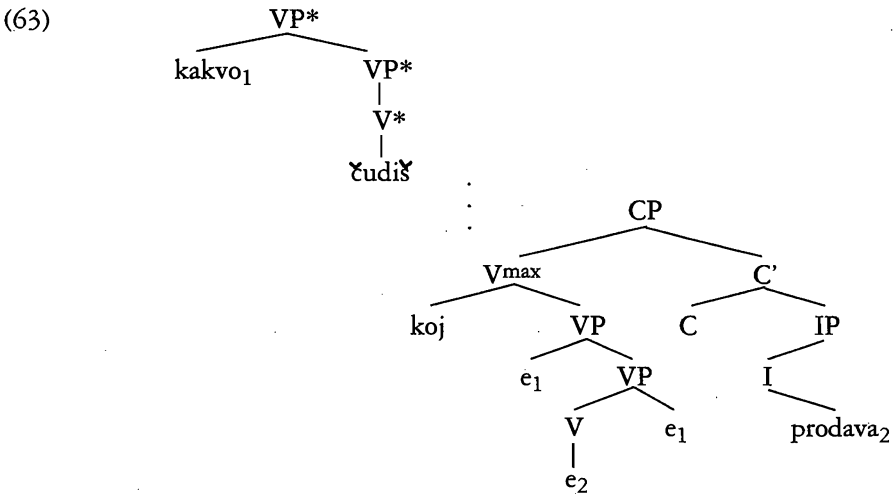


If the non-subcategorized wh-elements are not base-generated within  $V^{max}$ , the fact is derived that they can never appear in multiple questions, since these sentences exhibit a movement of  $V^{max}$  to the Spec or CP, a position which, once filled, does not accept further adjunctions (according to standard assumptions).

Rudin (p. 457) points out that, in Bulgarian, it is not possible for the wh-elements to move outside from a wh-island, as illustrated in (62):

- (62) *kakvo se čudivš koj znae koj prodava?*  
 what wonder-2sg who knows who sells  
 'What do you wonder who knows who sells?'

She claims that there are no syntactic reasons for this phenomenon. By contrast, our hypothesis leads to a configuration such as (63), where the object wh-element has been adjoined to the VP placed in the Spec of CP (an adjunction category) before adjoining to  $VP^*$ :



In (63), neither  $V^{max}$  nor CP are adjunction categories, and, furthermore,  $V^{max}$  occupies a non-thematic position, hence becoming an inherent barrier. On the other hand, CP becomes a barrier by inheritance. As a result, the movement of *kakvo* violates Subjacency. The same would happen if *koj* were moved. The only way of reaching the Spec of the matrix CP is by moving the whole  $V^{max}$ , as in (64) (the example is from Rudin 1988):

- (64) *koj kŭde misliš že B. iska da kažeš že ťte otide?*  
 who where think-2sg that B. wants t h a t say-2sg that will go-3sg  
 'Who do you think B. wants you to say will go where?'

5. Summary and Conclusions.

In this work we focussed on Slavic multiple questions. Rudin (1988) proposes that the underlying structure of multiple questions is not the same for all the Slavic languages. For Bulgarian (a [+MFS] language), the actual structure consists of a

multiple adjunction to the Spec of CP, and the others are adjoined to IP (8). Rudin considers this discrepancy to depend on the Condition on Spec-CP Adjunction (44). Her account predicts that the clitics have to appear to the right of the leftmost wh-element in the rest of Slavic languages. Rudin also proposes that the impossibility of simultaneously extracting several wh-elements from an embedded clause in these latter languages also follows by Subjacency.

First, we have found some flaws in the Subjacency-based account. We have seen that the concept of "barrier" as formulated in Chomsky (1986) is not compatible with Rudin's approach to multiple extraction, and, focussing on data from Slovak, we have proposed that it is Comp recursivity that accounts for that phenomenon if CP is considered as a "weak barrier" in these languages.

Furthermore, Comp recursivity is compatible with the behavior of Slovak clitics. Thus, the fact that just one constituent may appear between the complementizer and the clitics in embedded contexts is accounted for by claiming that, in these cases, there are two nodes C, the higher one for the complementizer and the lower one for the clitics. This account is also compatible with recent proposals on Scrambling as adjunction to IP.

Later, we have studied why it is obligatory for clitics to raise to C. We have explained this through several steps:

- a) in Slovak, TP is generated above AgrP;
- b) the node T<sup>-1</sup> subcategorizes for a slot which is the landing site for the head Agr (hence, we assume that head movement is substitution in these cases);
- c) taking pronominal clitics to be agreement elements, we claim that such a slot is the landing site for pronominal clitics as well;
- d) as is the case for the wh-operators, T has to raise at SS in Slovak to the same landing site where it raises at LF in other languages. We thus consider T to be an operator.

This facts force the clitics to always appear in a "second" position in the clause, once assumed that T has to raise to the highest C in the clause. This is a "second" position because it can only be preceded either by a constituent in the Spec of CP, or by a lexical verb which has raised to C. The fact that lexical verb raising is never compatible with the Spec of CP being already filled by a constituent is due to the combination of three factors:

- a) in languages where C has the features [+C, -I] (as in German, Slovak, etc), the presence of a constituent in the Spec of CP forces the presence of agreement morphemes in C;
- b) those agreement morphemes in C are weak in Pollock's (1989) sense, that is, a lexical verb cannot raise to a head where such morphemes are present, since this fact would give rise to a violation of the Theta-Criterion;
- c) in non-root clauses the lexical verb cannot appear in C either because, in these cases, expletive (and weak) agreement morphemes are also present.

Lastly, we have focussed on Bulgarian. We argued that the underlying structure for Bulgarian multiple questions is not (9), but (60b), where VP movement to the Spec of CP has taken place. This account offers several advantages:

- a) we avoid talking about adjunction to the right of the Spec of CP, a really odd movement in the GB-framework;
- b) the strict linear order between the wh-elements (subject/direct object/indirect object) is accounted for, since it is the order of base-generation of such constituents;
- c) we can also explain the fact that non-subcategorized wh-elements cannot appear in Bulgarian multiple questions if we consider such elements to be base-generated outside  $V_{max}$ ;
- d) a wh-element cannot be extracted from a wh-island in Bulgarian because  $V_{max}$  is an inherent barrier and CP is a barrier by inheritance;
- e) the whole group of wh-element can be extracted from an embedded clause, since this actually is  $V_{max}$  movement to the Spec of the matrix CP;
- f) lastly, we can also predict that Bulgarian clitics will appear to the right of the whole group of wh-elements, one of the salient differences between [+MFS] and [-MFS] languages.

## References

- Abney, S., 1987, *The English NP in its Sentential Aspect*. Doctoral Dissertation, MIT.
- Aoun, J. N., Hornstein, D. Lightfoot, and A. Weinberg, 1987, "Two Types of Locality", *LI*, 18, 537-577.
- den Besten, H. and G. Webelhuth, 1987, "Remnant Topicalization and the Constituent Structure of VP in the Germanic SOV languages", *GLOW Newsletter*, 18, 15-16.
- Borer, H., ed., 1986, *The Syntax of Pronominal Clitics*. "Syntax and Semantics", 19, Academic Press, Inc., Orlando, Florida.
- Chomsky, N., 1986, *Barriers*. MIT Press, Cambridge, Mass.
- Fanselow, G., 1989, "Scrambling as NP-movement", in Grewendorf and Sternefeld, eds.
- Grewendorf, G., and W. Sternefeld, eds., 1989, *Scrambling and Barriers*. John Benjamins. Amsterdam.
- Huang, C.-T. J., 1982, *Logical Relations in Chinese and the Theory of Grammar*. Doctoral Dissertation, MIT.
- Jaeggli, O. and K. Safir, eds., 1989, *The Null Subject Parameter*. Kluwer, Dordrecht.
- Kayne, R., 1989, "Null Subjects and Clitic Climbing", in O. Jaeggli and K. Safir, eds., 239-261.
- Koopman, H. and D. Sportiche, 1985, "Theta Theory and Extraction", *GLOW Newsletter*, 14, 57-58.
- Laka, I., 1989, "Constraints on Sentence Negation: The case of Basque", *MIT Working Papers*, 10, 199-216.
- Manzini, M. R. and K. Wexler, 1987, "Parameters, Binding theory, and Learnability", *LI*, 18, 413-444.
- Marácz, L., 1989, *Asymmetries in Hungarian*. Doctoral Dissertation, University of Groningen. [ASJU XXIV-2, -3 and XXV-1].
- Mascaró, J., and M. Nespó, eds., 1990, *Grammar in Progress*. Foris, Dordrecht.
- May, R., 1985, *Logical Form: Its Structure and Derivation*. MIT Press, Cambridge, Mass.
- Ortiz de Urbina, J., 1988, "Clausal Pied-piping", *Estudios de Lengua y Literatura*. Universidad de Deusto.
- Pesetsky, D., 1982, *Paths and Categories*. Doctoral Dissertation, MIT.
- , 1989, "Language-Particular Processes and the Earliness Principle", *GLOW Newsletter*, 22, 48-49.
- Plann, S., 1982, "Indirect Questions in Spanish", *LI*, 13, 297-312.

- Pollock, J.-V., 1985, "Sur la Syntaxe de *en* et le Paramètre du Sujet Nul", en M. Ronat and D. Couquaux, eds., 365-424.
- Raposo, E., 1986, "Some Asymmetries in the Binding Theory in Romance", *The Linguistic Review*, 5, 75-110.
- Rizzi, L., 1982, *Issues in Italian Syntax*. Foris, Dordrecht.
- , 1990b, "Speculations on Verb Second", in J. Mascaró and M. Nespó, eds., 375-386.
- , and I. Roberts, 1989, "Complex Inversion in French", *Probus*, 1.1, 1-30.
- Roberts, I., 1988, "Predicative APs", *LI*, 19, 703-710.
- Ronat, M. and D. Couquaux, eds., 1985, *La Grammaire Modulaire*. Minuit, Paris.
- Rudin, C., 1988, "On Multiple Questions and Multiple Wh-Fronting", *NLLT*, 6, 445-501.
- Santorini, B., 1989, "Against a Uniform Analysis of All Verb-Second Clauses", mss, University of Pennsylvania.
- Selkirk, E., 1982, *The Syntax of Words*. MIT Press, Cambridge, MA.
- Sportiche, D., 1988, "A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure", *LI*, 19, 425-449.
- , 1989, "Syntactic movement: constraints and parameters", *GLOW Newsletter*, 22, 62-63.
- Toman, J., 1986, "Cliticization from NPs in Czech and Comparable Phenomena in French and Italian", in H. Borer, ed., 123-148.
- Torrego, E., 1990, Spring lectures, Fundación Ortega y Gasset, Madrid.