

ON MULTIPLE QUESTIONS AND MULTIPLE  
WH FRONTING\*

0. INTRODUCTION

It is a familiar fact that languages differ in the extent to which they place Wh-words in a special, clause-initial position. Some languages, like English, normally place one and only one Wh-word in the specifier position of CP,<sup>1</sup> so in a multiple question like (1a), *what* is in SpecCP but *to whom* is in situ. Other languages, like Chinese, have all Wh-words in situ at S-structure, as in (1b). (In some languages, for instance, French, Wh-movement is optional; (1c) shows that French can follow either the English or the Chinese pattern.) A third group of languages, less well known to general linguists, but including all of the Slavic languages as well as some others, moves all Wh-words to a clause-initial position at S-structure, as in the Russian example in (1d), where *kto*, *čto*, and *kogda* are fronted.

- (1)a. **What** did you give **to whom**?
- b. Ni xiang-zhidao Lisi **weisheme** mai-le  
*you wonder Lisi why bought*  
**sheme?** (Huang 1982)  
*what*  
**What** do you wonder **why** Lisi bought (it)?
- c. **Qu'** as- tu donné à **qui**? = Tu as donne **quoi** à  
*what have you given to whom you have given what to*  
**qui?**  
*whom*

---

\* Early versions of most of the material in this article were presented at the LSA 1985 and 1986 annual meetings, the Fifth Balkan and South Slavic Meeting, and the Harvard Linguistic Circle. I would like to thank Dimitrije Stafanović, Ljuba Gribble, Rodića Bojoman, Rada Hanu, Vesna Radanović, Steven Franks, Eva Eckert, Donka Farkas, Donca Steriade, Milan Panić and several *NLLT* referees for helping with data and/or commenting on earlier drafts, and the University of Illinois Russian and East European Center for its support.

<sup>1</sup> Assuming the X-bar system of Chomsky (1986), in which S = IP (INFL phase), S' = CP (Comp phrase), and Wh-movement is to the specifier position of CP.

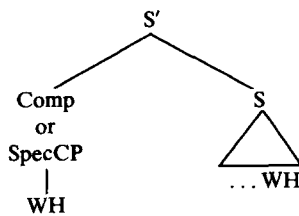
- (1d). **Kto čto kogda** skazal? (Wachowicz 1974)  
*who what when said*  
**Who said what when?**

It is this third group of languages, those with multiple fronting at S-structure, that I am primarily concerned with in this paper. The languages I will deal with in detail are Polish, Serbo-Croatian, Czech, Bulgarian (all slavic) and one Romance language, Romanian. Although multiple-Wh-fronting languages have been recognized in the generative linguistic literature at least since Wachowicz (1974), and several such languages have been discussed in the literature over the past decade, their structure has remained unclear. Analyses in which the Wh-words or phrases in questions like (1d) are all in some S-external position (Comp or Specifier of CP) have been put forward for example by Comorovski (1986) for Romanian, and by Rudin (1982, 1986) for Bulgarian. Other scholars have suggested that some or all of the Wh-words are in an S-internal position, perhaps adjoined to S. Reinhart (1982) argues for this view on universal grounds, as do Toman (1981) and Cichocki (1983) arguing from the facts of Czech and Polish. Lasnik and Saito (1984) also take this position; the language they consider is Polish. The two possible structures are shown schematically in (2c,d), with non-multiple-fronting structures in (2a,b) for comparison.

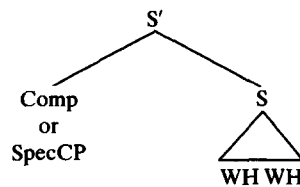
- (2) S-Structures:

non-multiple-fronting:

a. English

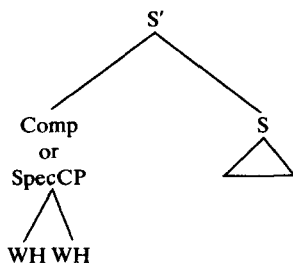


b. Chinese



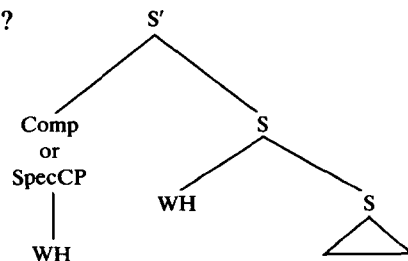
multiple Wh-fronting languages

c.



d.

or?



Analyses of multiple-fronting languages along the lines of (2d) were forced within early GOVERNMENT/BINDING THEORY (GB) by the fact that a multiply-filled Comp raises problems for standard GB notions such as the EMPTY CATEGORY PRINCIPLE (ECP) and the requirement that Wh-words c-command (and thus bind) their traces (see for instance Chomsky (1981, p. 52), Kayne (1981, p. 122)). These problems can however be overcome either by a revised definition of c-command (M-COMMAND (Aoun and Sportiche (1983), Chomsky (1986, p. 8); the m-command domain is the next highest maximal projection, not the next highest branching node, where Comp is not a maximal projection), or by devices such as Comp indexing which allow the Comp node itself to serve as the proper governor of at least one Wh-trace for purposes of satisfying the ECP (see e.g., Lasnik and Saito (1984); Aoun, Hornstein, Lightfoot and Weinberg (1987)). An indexing solution can easily extend to the analysis in which the landing site for Wh-movement is SpecCP rather than Comp, as will be demonstrated below. It is thus an open empirical question whether (2c) or (2d) is correct; nothing in current formulations of GB theory rules out either analysis. The major claim in this paper is that both (2c) and (2d) exist as S-structures, in different languages. Since I assume the phrase structure proposed by Chomsky (1986), my claim is more specifically that in some multiple Wh-fronting languages all Wh-phrases are in SpecCP at S-structure, while in others only one Wh-phrase is in SpecCP and the others occupy an IP-initial position; I will argue that they are adjoined to IP.

The view that all of the Wh-words in sentences like (1) are in SpecCP at S-structure in at least some languages is attractive partly because it parallels the structure that has recently been assumed for LOGICAL FORM (LF) by many linguists. Since Huang (1982) it has been widely accepted that Wh-movement applies at LF, not only in languages like Chinese, which have no syntactic Wh-movement, but also in languages like English when WH remains in situ in multiple questions. The LF structure of (3a) is thus (3b), in which the Wh-word which was unmoved at S-structure has been adjoined to SpecCP.

- (3)a. [<sub>SpecCP</sub> Who<sub>i</sub>] e<sub>i</sub> saw what?  
 b. [<sub>SpecCP</sub> What<sub>j</sub> [who<sub>i</sub>]] e<sub>i</sub> saw e<sub>j</sub>?

There is quite strong, although indirect and theory-dependent, evidence for structure (3b) on both semantic and syntactic grounds. Semantically, it accounts for the uniform interpretation of multiple questions, or what has been termed ABSORPTION (Higginbotham & May 1980; Aoun, Hornstein, and Sportiche 1980); that is, the fact that the Wh-words in a multiple question are interpreted as having the same scope. Syntactically,

LF movement leaving a trace subject to the ECP can account for the existence of SUPERIORITY phenomena in multiple questions.<sup>2</sup> The claim that some languages with multiple-Wh-fronting have all of the Wh-words in SpecCP at S-structure provides further concrete support for structures like (3b), and fills out a typology of Wh-movement in which languages choose freely at which level to move Wh-phrases to SpecCP. Some languages (e.g. Bulgarian, as we shall see) do all Wh-movement in the syntax; others (Chinese, for instance) do it all at LF, and still others (English) move just one WH in the syntax and any remaining ones at LF. All logical possibilities in (4) are then attested (French has the option of behaving like Chinese or like English):

(4)		<i>Chinese</i>	<i>French</i>	<i>English</i>	<i>Bulgarian</i>
	<i>single WH</i>	LF	LF/syntax	syntax	syntax
	<i>multiple WH</i>	LF	LF	LF	syntax

This not only provides support for the multiply-filled-Comp LF structure hypothesized as universal, but also makes available evidence for particular properties of this structure.

This article is organized as follows. In Section 1, I examine the evidence that both hypotheses mentioned above – that multiple Wh-words in sentences like those in (1c) are in Comp at S-structure, and that they are in S – are valid, in different languages. Although multiple-Wh-constructions in all multiple-fronting languages appear very similar in simple cases, on closer inspection they turn out to differ significantly in a number of ways. These differences cluster in a way which leads to the conclusion that there are two groups of multiple-fronting languages with quite different S-structures for multiple questions. Specifically, Bulgarian and Romanian have all of the Wh-words in SpecCP at S-structure (that is, they have a structure like (2c)), while Serbo-Croatian, Polish, and Czech have no more than one WH in SpecCP at S-structure (they have a structure like (2d)). For descriptive purposes and for ease of reference, I will label languages of the Bulgarian type as MULTIPLY-FILLED SPEC CP (+MFS) languages, while those of the Serbo-Croatian type are (–MFS) languages. In Section 2 I return to the issue of a crosslinguistic typology of Wh-movement and multiple-Wh-constructions. After discussing in more detail the internal structure of SpecCP and IP in the two types of multiple-Wh-fronting languages, I propose an account of the differences between them as well as among other, non-multiple-fronting types of

<sup>2</sup> See section 2 of this article. Contreras (1986) claims that it is also possible to account for the superiority facts at S-structure.

languages in terms of a parameterized CONDITION ON SPEC CP ADJUNCTION, based on Adams' (1984) Condition on Comp Adjunction. My condition disallows adjunction to SpecCP at specified levels of the grammar in a given language. Certain aspects of the structure of multiple questions, including in particular the optional or obligatory character of Wh-movement to SpecCP and/or IP adjunction in various languages, are not accounted for by the Condition on SpecCP Adjunction. These however are probably attributable to independent factors, such as the level at which Wh-movement occurs and the mechanisms available for marking a clause [+WH]. The article ends with a brief note on learnability.

### 1. TWO TYPES OF MULTIPLE-WH-FRONTING LANGUAGES

The five languages I will be most concerned with all have multiple Wh-fronting. The following, with all Wh-phrases at the beginning of the clause, are typical multiple questions.<sup>3</sup>

- (5)a. **Koj kogo** vižda? (Bulgarian)  
*who whom sees*  
**Who sees whom?**
- b. **Cine cu ce** merge? (Romanian)  
*who with what goes*  
**Who goes by what** (i.e. means of transportation)?
- c. **Ko koga** vidi? (Serbo-Croatian)  
*who whom sees*  
**Who sees whom?**
- d. **Kdo koho** videl? (Czech)  
*who whom saw*  
**Who saw whom?**
- e. **Kto co** robit? (Polish)  
*who what did*  
**Who did what?**

<sup>3</sup> For further examples and information on other Wh-constructions, including multiple free relatives, see Rudin (1986, 1988).

These simple multiple-Wh-questions look extremely similar. However, important dissimilarities among the five languages become evident when we turn to more complex multiple-Wh-constructions. Sections 1.1, 1.2, 1.3, and 1.4 present differences in multiple extraction, Wh-island effects, the position of clitics and parentheticals, and word order, respectively. In each case we will see that the behavior of multiple Wh-words in Bulgarian and Romanian indicates that they are all in SpecCP, while in Serbo-Croatian, Polish, and Czech only one of the fronted Wh-words is in SpecCP.

### 1.1. *Multiple Wh-extraction from a Clause*

The first difference between the two groups of languages that I will examine is the possibility for movement of multiple Wh-words out of an embedded clause. Such extraction is possible in the [+MFS] languages – Bulgarian and Romanian – but not in the [–MFS] type, which includes Serbo-Croatian, Polish, and Czech.

#### 1.1.1. *Bulgarian and Romanian*

In Bulgarian, all of the Wh-words in a multiple question must move up to the closest interrogative SpecCP, even if this means extracting more than one WH from a clause. Wh-words may not remain in situ, nor in the specifier position of a noninterrogative clause (2s = second person singular, 3p = third person plural, etc.)

- (6)a. **Koj kŭde** misliš [če e otišŭl \_\_ \_\_]?  
*who where think-2s that has gone*  
**Who** do you think (that) went **where**?
- b. \***Koj** misliš [če e otišŭl \_\_ **kŭde**]?  
*who think-2s that has gone where*
- c. \***Kŭde** misliš [če **koj** e otišŭl \_\_]?  
*where think-2s that who has gone*
- d. \***Koj** misliš [**kŭde** (če) e otišŭl \_\_ \_\_]?  
*who think-2s where that has gone*
- e. \***Kŭde** misliš [**koj** (če) e otišŭl \_\_ \_\_]?  
*where think-2s who that has gone*

This is true regardless of the type of Wh-phrases involved. Examples

with object Wh-phrases are given in (7). (**Boris** in these sentences is a Topic.<sup>4</sup>)

- (7)a. **Boris na kogo kakvo** kaza [če šte dade \_\_ \_\_]?  
*Boris to whom what said that will give-3s*  
**What** did Boris say that (he) would give **to whom**?
- b. \***Boris na kogo** kaza [če šte dade **kakvo** \_\_]?  
*Boris to whom said that will give what*
- c. \***Boris kakvo** kaza [če šte dade \_\_ **na kogo**]?  
*Boris what said that will give to whom*
- d. \***Boris kakvo** kaza [**na kogo** šte dade \_\_ \_\_]?  
*Boris what said to whom will give-3s*
- e. \***Boris na kogo** kaza [**kakvo** šte dade \_\_ \_\_]?<sup>5</sup>  
*Boris to whom said what will give-3s*

However, this is not to say that multiple Wh-words can always be extracted from a clause. For instance, it is (not surprisingly) impossible to front both a relative and an interrogative WH to the same higher clause. As (8a–b) show, an interrogative clause can contain only interrogative Wh-words (or complementizer) in SpecCP, and a relative clause only relative Wh-words (or the relative complementizer **deto** in Comp). However, it is possible to extract both interrogative and relative Wh-phrases from a clause if each WH is able to move to an appropriate higher SpecCP position, as in (8c).

- (8)a. \***Čoveka, kojto kakvo** ne znaeš [(če) e  
*the man who-REL what NEG know-2s that has*  
 kupil \_\_ \_\_] . . .  
*bought*

<sup>4</sup> Topic position is pre-SpecCP, presumably adjoined to CP, and a topic NP can occur in any type of clause, with anything in SpecCP and/or Comp. See Rudin (1986, ch. 2) for further details.

<sup>5</sup> This sentence is OK with the reading where **na kogo** is the object of **kaza** 'said', i.e. where it originates in the upper clause. It cannot be taken as the object of **dade** 'give', as indicated by the blank in the example.

(8)b. \*Čoveka, (deto) ne znaeš [kakvo kojto e  
*he man that-REL NEG know-2s what who-REL has*  
 kupil \_ \_]...  
*bought*

c. Čoveka, kojto ne znaeš [kakvo kazvat [če e  
*the man who-REL NEG know-2s what say-3p that has*  
 kupil \_ \_]]...  
*bought*

The man **who** you don't know **what** they say that he bought...

Aside from this and a few other exceptions, extraction of multiple Wh-words is grammatical in Bulgarian. In Romanian as well, multiple Wh-movement out of a clause is possible, as has been noted by Adams (1984) and Comorovski (1986). To give just one example, all Wh-words must be fronted to the matrix clause SpecCP in questions like (9), where the subordinate clause is noninterrogative. This sentence with one or more of the WHs left behind, either in situ or in the lower SpecCP position, would be ungrammatical.

(9)a. Cine cui ce ziceai [că i -a  
*who to whom what said-2s that to him has*  
 promis \_ \_ \_]?<sup>6</sup>  
*promised*

**Who** did you say promised **what to whom**?

Even Wh-words that originate in different clauses can move to a single higher SpecCP in Romanian:<sup>7</sup>

(10) Cine ce ziceai [că \_ își închipuie [că ai  
*who what said-2s that to himself imagines that have-2s*  
 descoperit \_ \_]]?  
*discovered*

**Who** did you say imagines you discovered **what**?

<sup>6</sup> Examples (9) and (10) are from Comorovski (1986).

<sup>7</sup> This does not seem to be true in Bulgarian. For some discussion of differences between Bulgarian and Romanian, see Kraskow (1987).



Thus, in both Bulgarian and Romanian movement of more than one Wh-word out of the clause in which they originate is fully normal and often obligatory.

1.1.2. *Serbo-Croatian, Polish, and Czech*

In the [-MFS] languages, on the other hand, multiple Wh-extraction from a clause is ungrammatical. In this section I briefly discuss the situation in Serbo-Croatian, Polish, and Czech in turn.

In Serbo-Croatian multiple questions involving movement out of a clause, all but one of the Wh-words can and for most speakers<sup>8</sup> must remain in situ. Extraction of a second (or third) Wh-word from the clause, as in (11c,d), is ungrammatical, as is movement to the SpecCP position of the lower clause (11e,f). Note that the in situ Wh-word is however normally placed to the left of the verb, the position in which focused material occurs. (Examples from Steven Franks.)

- (11)a. **Ko** želite [da vam **šta** kupi \_\_\_]?  
*who want-2p to you what buy-3s*  
**Who** do you want to buy you **what**?
- b. **Šta** želite [da vam **ko** kupi \_\_\_]?  
*what want-2p to you who buy-3s*  
**What** do you want **who** to buy you?

---

<sup>8</sup> Some speakers do accept sentences like (i), with multiple Wh-words moved out of their clause, but even these speakers also accept the same sentences with one WH left behind, as in (ii).

(i) **Ko koga** misliš [da je video \_\_\_ \_\_\_]?  
*who whom think-2s that has seen*

(ii) **Ko** misliš [da je **koga** video \_\_\_]?/**Koga** misliš [da je **ko** video \_\_\_]?

Furthermore, **koga** in (i) is not in SpecCP; clitics are placed before it (see section 1.3):

(iii) \***[ko koga]** si mislio [da je video \_\_\_ \_\_\_]  
*who whom have thought-2s that has seen-3s*

(iv) **[ko]** si **koga** mislio [da je video \_\_\_ \_\_\_]  
*who have-2s whom thought that has seen-3s*

**Who** did you think had seen **whom**?

It seems that for such speakers focusing as well as Wh-movement can operate across clause boundaries, that is, they allow cyclic adjunction to successive IP as well as SpecCP nodes, perhaps only with certain bridge verbs.

- (11)c. \***Ko** **šta** **želite** [da vam kupi \_\_ \_\_]?  
*who what want-2p to you buy*
- d. \***Šta** **ko** **želite** [da vam kupi \_\_ \_\_]?  
*what who want-2p to you buy-3s*
- e. \***Ko** **želite** [**šta** da vam kupi \_\_ \_\_]?  
*who want-2p what to you buy*
- f. \***Šta** **želite** [**ko** da vam kupi \_\_ \_\_]?  
*what want-2p who to you buy*

In Polish, as in Serbo-Croatian, long movement of multiple Wh-words is impossible. It is more difficult to test this in Polish, since any extraction at all from a finite clause is normally ungrammatical. However, subjunctive complements do allow Wh-movement across a clause boundary. Example (12a), from Lasnik and Saito (1984), is grammatical. But an otherwise identical sentence with two Wh-words fronted out of the clause is not:

- (12)a. **Co** Maria chce, żeby Janek kupił?  
*what Maria wants that Janek buy*  
**What** does Maria want Janek to buy?
- b. \***Co** **komu** Maria chce, żeby Janek kupił?  
*what to whom Maria wants that Janek buy*  
**What** does Maria want Janek to buy **for whom**?

Certain BRIDGE verbs also permit single Wh-words to be extracted, as in (13a), from Cichocki (1983), but here again multiple Wh-words may not be extracted.

- (13)a. **Co** Janek mówił, że studenci czytają?  
*what Janek said that students read*  
**What** did Janek say that the students read?
- b. \***Dlaczego** **co** Janek mówił, że studenci czytają?  
*why what Janek said that students read*  
**What** did Janek say the students read **why**?

Like the other [-MFS] languages, Czech also allows extraction of only one Wh-word from a clause. Long distance movement of a single Wh-word is fully grammatical (examples from Toman (1981)):

- (14)a. **Kde** si myslíš, že budeme spát?  
*where REFL think-2s that will-1p sleep*  
**Where** do you think that we will sleep?

But in spite of this, extraction of more than one WH from a clause is impossible:

- (15)a. \***Kde kdy** si myslíš, že budeme spát?<sup>9</sup>  
*where when REFL think-2s that will-1p sleep*  
**Where** do you think we will sleep **when**?
- b. \***Na koho kdy** si myslíš, že čekal?  
*For whom when REFL think-2s that waited-3s*  
**Who** do you think that he was waiting for **when**?

### 1.1.3. Multiple Extraction and the MFS Hypothesis

In the last two sections we have seen that overt syntactic movement of more than one Wh-element to a higher clause, resulting in a configuration like (16), is possible in Bulgarian and Romanian but not Serbo-Croatian, Polish, or Czech.

- (16) [CP WH<sub>i</sub> WH<sub>j</sub> ... [CP ... e<sub>i</sub> ... e<sub>j</sub> ...]]

Given standard GB assumptions, this fact supports the contention that Bulgarian and Romanian allow multiple WHs in (adjoined to) SpecCP in the syntax, i.e. at or before S-structure, while the other languages do not permit multiple WHs to adjoin to SpecCP except at LF.<sup>10</sup> In order to produce a structure like (16) without violating subadjacency it is necessary for more than one Wh-phrase to pass through the embedded clause specifier position: a more detailed structure for (16), showing the traces in the lower SpecCP, is (17).

- (17) [CP WH<sub>i</sub> WH<sub>j</sub> ... [CP [SpecCP t<sub>i</sub> [t<sub>j</sub>]] ... e<sub>i</sub> ... e<sub>j</sub> ...]]

Thus Bulgarian and Romanian must allow multiple Wh-traces to be adjoined to SpecCP at S-structure in order to have S-structures like (16).

It might be objected that the Wh-traces in (17) are not necessary if IP

<sup>9</sup> Eva Eckert (p.c.) reports that **Kde si kdy myslíš, že budeme spát** 'Where do you think (that) we will sleep when?', with the clitic *si* between the two Wh-words, is possible for her in colloquial speech; this is parallel to the marginal Serbo-Croatian facts discussed in note 8.

<sup>10</sup> I assume that Wh-movement of a second or subsequent WH is adjunction to SpecCP. This assumption is discussed in section 2.

is not a bounding node for subjacency. However, Comorovski (1986) presents independent evidence based on a V-fronting rule that the bounding node in Romanian is IP, not CP. Furthermore, neither Bulgarian nor Romanian has any restriction on how many clauses the Wh-words may cross; even if CP were the bounding node in these languages, the Wh-phrases would have to pass through the SpecCP positions of all but the lowest clause. Sentence (10) above shows multiple Wh-movement across more than one CP boundary in Romanian, and (18) demonstrates that this is possible in Bulgarian as well.

- (18) **Koj kŭde** misliš [CP *če* Boris iska [CP *da* kažeš [CP  
*who where think-2s that Boris wants to say-2s*  
*če šte otide — —]]]?  
*that will go-3s*  
**Who** do you think Boris wants you to say will go **where?**  
 (Bulgarian)*

The [−MFS] languages, which do not allow the surface syntactic configuration in (16), are prevented from allowing it under the analysis suggested here by the fact that they are not able to have multiple Wh-traces adjoined to SpecCP at S-structure.

### 1.2. *Wh-Islands*

Very closely related to the issue of multiple Wh-extraction to a higher clause is that of the island status of embedded questions. As Comorovski (1986) has noted, given subjacency as the explanation of Wh-islands, we predict that a language that allows multiple Wh-elements in Comp at the level at which Wh-movement occurs “will not obey any form of the Wh-island Constraint”, since in such a language a Wh-phrase could not be blocked from moving through or leaving a trace in a Comp that contains another WH. Adams (1984) makes the same observation, and of course it holds equally well if we substitute SpecCP for Comp as the Wh-position. In the present case, we predict that Bulgarian and Romanian will not have Wh-islands, but Serbo-Croatian, Polish, and Czech will, given our hypothesis that Bulgarian and Romanian are [+MFS] while the other three languages are [−MFS]. This prediction holds, as we shall now see.

Bulgarian freely allows extraction of Wh-words from an embedded question, or even from several interrogative Wh-clauses, as in (19).

- (19) Vidjah edna kniga, **kojato**<sub>i</sub> se čudja [**koj** znae [**koj**  
*saw-1s a book which wonder-1s who knows who*  
 prodava <sub>i</sub>]]  
*sells*  
 I saw a book **which** I wonder **who** knows **who** sells (it).

Movement of an interrogative Wh-word out of an embedded Wh-question is usually felt to be less normal than movement of a relative pronoun, but is fairly acceptable with heavy Wh-phrases, as in (20b).<sup>11</sup>

- (20)a. \***Kakvo** se čudiš **koj** znae **koj** prodava?  
*what wonder-2s who knows who sells*  
**What** do you wonder **who** knows **who** sells?
- b. ?**Koja ot tezi knigi** se čudiš **koj** znae **koj** prodava?  
*which of these books wonder-2s who knows who sells*  
**Which of these books** do you wonder **who** knows **who** sells?

This asymmetry between questioning and relativization is similar to the situation in some Scandinavian languages, where it has been suggested that semantic or pragmatic constraints make it more difficult to question than to relativize out of a question (Lie 1982). The difficulty of questioning out of a question extends to embedded yes-no questions as well: As the following examples show, questioning out of a noninterrogative clause and relativizing out of a question are fine, but questioning out of a question produces a sentence of quite low acceptability. Thus is not simply a Wh-word in SpecCP that tends to block extraction in sentences like (20), but some other, perhaps nonsyntactic, factor.

- (21)a. **Koj** misliš če e došul?  
*who think-2s that has come*  
**Who** do you think has come?
- b. ... čoveka, **kojto** se čudiš dali e došul ...  
*the person who wonder-2s whether has come*  
 ... the person **who** you wonder whether he has come ...

<sup>11</sup> Even better is a variant with a preposition: **Za koja ot tezi knigi se čudiš koj znae koj prodava** 'For which of these books, you wonder who knows who sells'. However, it is not clear to me that the phrase 'for which of these books' has undergone Wh-movement. Rizzi (1982) notes a similar effect of 'heaviness' of the extracted Wh-phrase in Italian as well.

- (21)c. ??**Koj** se čudiš dali e došul?  
*who wonder-2s whether has come*  
**Who** do you wonder whether came?

Romanian allows Wh-movement of both interrogative (22a) and relative (22b) Wh-phrases from inside an indirect question, although, as in Bulgarian, it is more difficult to construct acceptable examples with interrogative Wh-movement. There is no limit on the number of clauses that may intervene between the fronted Wh-phrase and the extraction site (examples from Comorovski (1986)).

- (22)a. **Pentru care clauză<sub>i</sub>** vrei să afli cine nu a  
*for which paragraph want-2s to learn who not has*  
 decis încă ce va vota \_\_\_i?  
*decided yet what will-3s vote*  
**For which paragraph** do you want to learn **who** has not  
 decided yet **what** he will vote?
- b. **Maria, cu care<sub>i</sub>** știu ce crezi că am  
*Maria with whom know-1s what think-2s that have-1s*  
 discutat \_\_\_i, ...  
*discussed*  
**Maria, with whom** I know **what** you think I discussed, ...

In short, as expected given the MFS hypothesis, Bulgarian and Romanian, which are [+MFS] languages, do not respect Wh-islands. Furthermore, the lack of Wh-island effects is not due to an absence of subjacency effects in general in these languages. Both Bulgarian and Romanian do obey other subjacency islands; for example, they do not allow movement from inside a Complex NP.

- (23)a. \*Tova e momčeto **na koeto** misulța ce (mu)  
*this is the boy to whom the thought that to him*  
 dadohme bonboni jadosva lekara. (Bulgarian)  
*gave-1p candy angers the doctor*  
**This is the boy to whom** the thought that we gave (him) candy  
 makes the doctor angry.

- (23)b. \*Ambasadorul, **pe care** nu ați crezut zvonul că  
*the ambassador whom not have-2s believed the rumor that*  
 teroriști l- au răpit, nu a mai  
*terrorists him have kidnapped not has more*  
 fost văzut de trei zile. (Romanian)  
*been seen for three days*

The ambassador, **whom** you didn't believe the rumor that the terrorists had kidnapped, has not been seen for three days. (Comorovski 1986)

Turning now to the [-MFS] languages, we see that the situation is reversed; Wh-islands do in general hold. In Serbo-Croatian, neither questioning nor relativization is possible from inside a Wh-question, regardless of whether it is a subject or nonsubject Wh-phrase that is extracted and of whether a resumptive pronoun appears:

- (24)a. \***Šta** si me pitao **ko** može da uradi?  
*what have-2s me asked who can to do*  
**What** did you ask me **who** can do?
- b. \*... osoba, **koja** sam ti rekao **gde** (on) živi ...  
*individual who have-1s you told where he lives*  
 ... the individual **who** you asked me **where** (he) lives ...

Polish also does not allow either interrogative or relative Wh-movement from inside a finite Wh-clause<sup>12</sup> (examples (25a,b) are from J. R. Williams (1968b)).

- (25)a. \***Co** on zapytał [<sub>S</sub> **кто** wynalazł \_\_\_]?  
*what he asked who invented*  
**What** did he ask **who** invented?
- (25)b. \*... maszyna **która** on zapytał [<sub>S</sub> **кто** wynalazł \_\_\_] ...  
*machine which he asked who invented*  
 ... the machine **which** he asked **who** invented ...

Cichocki (1983) points out that relativization from an infinitival Wh-clause, as in (26), is possible. However, he argues convincingly, on the

<sup>12</sup> In fact, as we have seen Polish generally does not permit Wh-extraction from a finite clause in any case.

basis of facts involving conjunction, that the Wh-word in the embedded clause is not in SpecCP (in his terminology, it is not in COMP<sub>1</sub>)<sup>13</sup> but rather in the position occupied by Wh-words other than the first one in a multiple question. That is, (26a) has the structure shown in (26b), and does not involve a Wh-word and a trace together in specifier position.

- (26)a. Pytanie **na które** zdecydowałeś się [jak odpowiedzieć]  
*question on which decided-2s REFL how to answer*  
 było trudne.  
*was difficult*

The question **which** you decided **how** to answer was difficult.

- b. pytanie [<sub>CP</sub> [<sub>Spec</sub> **na które**<sub>i</sub>] . . . [<sub>CP</sub> [<sub>Spec</sub> **t<sub>i</sub>**] **jak** opowiedzieć **t<sub>i</sub>**] . . .

In Czech, some though not all speakers do accept apparent Wh-island violations like (27a) and (27b) ((\* indicates idiolectal variability):

- (27)a. (\*)**Kdo** se tě ptal **co** dělá?  
*who have-3p you asked what does*  
**Who** did they ask you **what** (he) does?

- b. To je ten, **o kom** jsem ti řekl, **co** dělá.  
*that is the one about whom have-1s you told what does*  
 That is the person **who** I asked you **what** (he) does.

- c. \*To je ten, **kdo** jsem ti řekl, **co** dělá.  
*that is the one who have-1s you told what does*

However, note that (27b) is not a Wh-island violation at all: the oblique case form **o kom** indicates that this Wh-phrase is the object of **řekl** and originated in the same clause it occupies at S-structure. The nominative **kdo**, the expected form for the D-structure subject of the most deeply embedded clause, is impossible, as shown in (27c). Sentence (27a), for those speakers who accept it, is more problematic, but still not a major problem for the [-MFS] hypothesis. It may be possible to explain it along the same lines as the Polish island violation in (26). And even if this turns out to be wrong, it is still possible within the account of multiple fronting

<sup>13</sup> Cichocki's two-Comp analysis of Polish is discussed in Section 2.3 below.



constructions proposed in the second part of this paper to account for this variety of Czech in terms of a condition on Phonetic Form.

To summarize this section, the Wh-island data, like the multiple extraction facts, support the contention that the multiple Wh-fronting languages fall into two groups, differing in the position of fronted Wh-words and the structure of SpecCP at S-structure. Bulgarian and Romanian, which I analyze as allowing multiple Wh-phrases in SpecCP at S-structure, show exactly the predicted lack of Wh-island effects. On the other hand, Serbo-Croatian, Polish, and to a large extent also Czech do exhibit Wh-island effects, just as expected given the hypothesis about [-MFS] languages.

### 1.3. Evidence for Constituent Structure

#### 1.3.1. Clitic Position

Another respect in which multiple Wh-constructions in the [+MFS] and [-MFS] languages differ is the position of clitics relative to the Wh-words. Clitic position in the [-MFS] languages is most often between the first Wh-phrase and the second one, although there are sometimes other options too. In the [+MFS] languages, clitics are placed after the whole series of Wh-phrases. In this section I demonstrate that this difference in clitic position reflects the different constituent structure of the fronted Wh-word group in the two sets of languages. I begin by comparing Bulgarian and Serbo-Croatian and then look at the remaining languages.

In Bulgarian, as shown in (28) and (29), clitics (in this case the auxiliary *e* and pronoun *ti*) may not split up the Wh-word sequence. In Serbo-Croatian, on the other hand, the clitic pronouns and auxiliaries must directly follow the first WH-word, as can be seen in (30) and (31). No other position is possible; in particular clitics may not come at the end of the sequence of Wh-words.

(28)a. **Koj kakvo ti e kazal?** (Bulgarian)

*who what you has told*

**Who told you what?**

b. **\*Koj ti e kakvo kazal?**

*who you has what told*

(29)a. **Koj kakvo na kogo e dal?** (Bulgarian)

*who what to whom has given*

**Who gave what to whom?**

- (29)b. \***Koj kakvo e na kogo dal?**  
*who what has to whom given*
- c. \***Koj e kakvo na kogo dal?**  
*who has what to whom given*
- (30)a. **Ko mu je šta dao?** (Serbo-Croatian)  
*who him has what given*  
**Who gave him what?**
- b. \***Ko šta mu je dao?**  
*who what him has given*
- (31)a. **Ko je što kome dao.** (Serbo-Croatian)  
*who has what to whom given*  
**Who gave what to whom?**
- b. \***Ko što je kome dao?**  
*who what has to whom given*
- c. \***Ko što kome je dao?**  
*who what to whom has given*

It might be objected that these facts could reflect a difference not in the position of Wh-words but rather in the clitic placement rules of Bulgarian and Serbo-Croatian. And in fact cliticization does work differently in the two languages: in Serbo-Croatian, all clitics are clause-second, while in Bulgarian they are proclitic to the verb. However, the clitic placement facts do nonetheless provide evidence for a difference in the constituent structure of the Wh-word sequence in the two languages.

SECOND POSITION clitics in Serbo-Croatian are placed either after the first word or after the first major constituent of the sentence. For instance, in (32a-b) (Browne 1976) the clitic **je** can follow either a full NP or the first word of the NP:

- (32)a. Njegova prijateljica **je** profesor književnosti.  
*his friend-FEM is professor of literature*  
 His friend is a professor of literature.
- b. Njegova **je** prijateljica profesor književnosti.  
*his is friend-FEM professor of literature*

In (30a) and (31a) the clitics can follow **ko** simply because it is the first

word of the sentence, regardless of constituent structure. But (30b) and (31b, c) do not give conclusive evidence concerning constituency. The fact that clitics cannot follow a clause-initial sequence of Wh-words such as **ko šta** or **ko šta kome** proves decisively that such a sequence is not a constituent in Serbo-Croatian, unlike in Bulgarian, where, as we shall see below, there is good evidence that a sequence of fronted Wh-words such as **koj kakvo** or **koj kakvo na kogo** is a constituent.

A *NLLT* reviewer has suggested that the data in (28)–(31) could be due to a verb movement rule operating in Serbo-Croatian but not Bulgarian; pronominal clitics in Serbo-Croatian would front along with the verb to which they are attached. This is clearly not the case, however, since clitic placement is the same in tenses with no cliticizing auxiliary verb, and verbs other than a very limited set of auxiliaries do not front (REFL CL-reflexive clitic):

- (33)a. **Koliko** im **ko** daje?  
*how much to them who gives*  
**Who** gives them **how much**?
- b. **Ko** se **čega** boji?  
*who REFL CL of what is afraid*  
**Who** is afraid of **what**?

I will not go into further detail here, but there is a great deal of evidence that clitics in Serbo-Croatian attach to the first element (word or constituent) in the clause, not to the verb. The first element specifically includes any filled Comp or SpecCP; see for instance Browne (1976, 1986, p. 25).

What are the predictions for clitic position in the other multiple-Wh-fronting languages? The answer obviously depends on the specific clitic placement rules of each language, but some general predictions can be made. In particular, in the case where clitic placement is by encliticization to the first constituent of the clause (or by movement to Comp), we should find that clitics follow the whole group of fronted Wh-words if the WHs form a constituent (i.e. in the [+MFS] languages, where they are all in SpecCP), but that they occur after the first one of the Wh-words when it does not form a constituent with the second and subsequent WHs (i.e. in the [–MFS] languages):

- (34)a. [<sub>SpecCP</sub> WH WH WH] clitics [<sub>IP</sub> . . .] ([+MFS] languages)  
 b. [<sub>SpecCP</sub> WH] clitics [<sub>IP</sub> WH WH . . .] ([–MFS] languages)

Thus Romanian should place clitics at the end of the Wh-word sequence, and Polish and Czech should put them after the first Wh-word – again, if their clitic position is after the first constituent. And indeed, this is essentially the pattern we find.

In Romanian, unfortunately, clitics are not clause-second, but rather proclitic to V. The clitics, including clitic pronouns like *le* ‘them’ in (35) as well as the auxiliary *a* ‘has’, do follow the entire Wh-word sequence rather than just the first Wh-word:

- (35)a. Nu ştiu **cine ce** le- a spus.  
*NEG know-1s who what them has told*  
 I don't know **who** told them **what**.
- b. \*Nu ştiu **cine** le- a **ce** spus.  
*NEG know-1s who them has what told*

But of course this proves nothing about the constituent status of the Wh-words, since the ungrammaticality of (35b) is due to the clitics' being separated from the verb. Nonetheless, these facts do not contradict the pattern in (34), and are thus at least compatible with the [+MFS] hypothesis.

Both Polish and Czech do observe some version of Wackernagel's law, and from them we get more direct evidence for structure (34b). In Polish, clitics have two options: they can appear in clause-second or preverbal position. But unlike in Serbo-Croatian, CLAUSE-SECOND POSITION always means following the first major constituent. A clitic is never placed after the first word of a sentence if that word is part of a nonclausal maximal projection, such as NP (see (36)) – or SpecCP. The following examples are from Toman (1981).

- (36)a. [<sub>NP</sub> Ten stary pan] wczoraj **się** ogolił.  
*this old man yesterday REFL shaved-3S*  
 This old man shaved **himself** yesterday.
- b. [<sub>NP</sub> Ten stary pan] **się** wczoraj ogolił.  
*this old man REFL yesterday shaved*
- c. \*[<sub>NP</sub> Ten **się** stary pan] wczoraj ogolił.  
*his REFL old man yesterday shaved*

Clitics can however follow the first of two Wh-words:

(37)a. **Kto się komu** podoba?

*who REFL to whom likes*

**Who likes who?**

b. **Kto komu się** podoba?

*who to whom REFL likes*

Sentence (37b) represents the option of placing a clitic (*się*, in this case) adjacent to the verb regardless of what precedes (as in (36a)); it gives no information about the constituent status of the string of Wh-words. However, (37a) does give such information. Since second-position clitics in Polish must follow the first major constituent, not the first word, a Wh-sequence like **kto komu** must not be a constituent. Cichocki (1983) gives further evidence for this point. In triple questions, clitics may follow the first Wh-word or may appear in the preverbal position at the end of the whole Wh-word group, but they may not follow the second Wh-word. The possibilities for placement of the clitic *by* in (38) indicate that the first Wh-word alone comprises a major constituent (SpecCP), but the first two together do not.

(38)a. WH (clitics) WH (\*clitics) WH (clitics) V . . .

b. **Kto by komu jaką** napisał książkę?

*who would to whom what kind write book*

**Who would write what kind of book for whom?**

c. \***Kto komu by jaką** napisał książkę?

*who to whom would what kind write book*

d. **Kto komu jaką** by napisał książkę?

*who to whom what kind would write book*

The clitic situation is simpler in Czech than in the other [–MFS] languages; clitics always follow the first major constituent of the clause. In the following example, adapted from Toman (1986), the only grammatical position for the clitics *se* and *ti* is immediately after the sentence-initial NP; in particular, the presence of the adverb **jednou** makes it clear that the clitics may not follow more than one constituent (39c), nor may they interrupt a constituent (39b):

(39)a. Tohle staré kolo **se ti** jednou rozpadne.

*this old bicycle REFL to you once falls apart*

This old bicycle will fall apart on you one day.

- (39)b. \*Tohle **se ti** staré kolo jednou rozpadne.  
 c. \*Tohle staré kolo jednou **se ti** rozpadne.  
 d. \*Tohle staré kolo jednou rozpadne **se ti**.

Under the [-MFS] analysis, in which the first Wh-word in a multiple question is set off by itself in SpecCP, clitics in multiple Wh-constructions would be expected always to follow the first Wh-word in Czech. None of the other options available in Serbo-Croatian and Polish should be available. And exactly this is the case (examples are from Toman (1981 and 1986)):

- (40)a. **Kdo ho kde** videl je nejasné.  
*who him where saw is unclear*  
**Who** saw him **where** is unclear.
- b. \***Kdo kde** ho videl je nejasné.  
*who where him saw is unclear*
- (41)a. **Komu ho kdy** dali, nikdo neví.  
*to whom him when gave-3p nobody knows*  
 Nobody knows **when** they gave him **to whom**.
- b. \***Komu kdy** ho dali, nikdo neví  
*to whom when him gave-3p nobody knows*

In summary, the position of clitics provides good evidence in each of the [-MFS] languages studied that the fronted Wh-words in multiple questions do not form a constituent. Although the details of clitic placement are different in Serbo-Croatian, Polish, and Czech, in all three languages clitics which are placed after the first constituent of the clause appear immediately following the *first* of the fronted Wh-phrases. In no case can clitics appear after the second of three Wh-phrases, and in the one case where they can appear at the end of the entire Wh-word sequence, in Polish, it is because the language independently allows the option of procliticization to V. These facts clearly support an analysis in which the first Wh-word alone is a constituent separate from the other WHs: the first WH is in SpecCP, while the others are not. The clitic-position evidence for constituent structure of the Wh-word group is inconclusive for the [+MFS] languages, unfortunately, since in both Bulgarian and Romanian clitics are pre-V, not clause-second. However, clitics do not provide any arguments *against* the hypothesis that all of the fronted Wh-words are in SpecCP in these languages, and we will see in

the next section that there are independent reasons to consider the sequence of Wh-phrases in [+MFS] languages to be a single constituent.

### 1.3.2. *Parentheticals and Related Phenomena*

Another source of evidence for a difference in constituent structure between the two groups of languages is the possibility or impossibility of interrupting the Wh-word sequence with adverbs, particles, or parenthetical phrases. Once again, as with each area we have looked at, Romanian and Bulgarian pattern together, while Polish and Czech are like Serbo-Croatian. In the [+MFS] languages, generally nothing may intervene between the fronted Wh-words, but in the [-MFS] languages various types of material can appear after the first Wh-word, further supporting the idea that the first WH is a separate constituent in the [-MFS] but not in the [+MFS] languages.

A typical example of the difference in freedom of occurrence of adverbs is the contrast between (42) and (43). In the Bulgarian sentence the adverb meaning 'first' (*prv*) cannot be placed between the two Wh-words. But in the parallel Serbo-Croatian example, the adverb may appear between them. (Recall that the clitic *je* must be in second position in the clause; *prvi* therefore cannot come immediately after *ko*.) Although some speakers find (43a) slightly more natural than (43b), (43b) is far more normal than the totally ungrammatical (42b) in Bulgarian.

(42)a. Zavisi ot tova, **koj kogo prv** e udaril. (Bulgarian)  
*depends on this who whom first has hit*

It depends on **who** hit **whom** first.

b. \*Zavisi ot tova, **koj prv kogo** e udaril.  
*depends on this who first whom has hit*

(43)a. **Ko je koga prvi** udario? (Serbo-Croatian)  
*who has whom first hit*

**Who** hit **whom** first?

b. **Ko je prvi koga** udario?  
*who has first whom hit*

I do not have complete data on the possibility of adverbs appearing within the Wh-word group in the other multiple-fronting languages, but the contrast between Bulgarian and Serbo-Croatian is certainly suggestive, and would be predicted to extend to the other languages as well.

Parenthetical phrases, like sentence adverbs, tend to occur most comfortably at major constituent boundaries. In view of this fact, it is of interest to investigate which positions parenthetical phrases may occupy in multiple questions. In Romanian and Bulgarian, parentheticals do not easily split the fronted Wh-word sequence. Comorovski (1986) states that no parentheticals may intervene between the Wh-phrases in a multiple question in Romanian, but in fact the data on this point are somewhat murky. Speakers I have consulted give unclear and sometimes contradictory judgements of sentences like those in (44); the “?” markings represent the fact that speakers are hesitant about accepting either variant.

(44)a. ?\***Cine**, după părerea ta, **ce** a făcut?  
*who after opinion your what has done*  
**Who**, in your opinion, did **what**?

b. ?\***Cine ce**, după părerea ta, a făcut?  
*who what after opinion your has done*

Although the awkwardness of (44b) is puzzling, the low acceptability of (44a) is just what we would predict, given that **cine** and **ce** form a constituent (cf. (32a)). The problem for speakers with the examples in (44) is not simply that they are unwilling to accept the parenthetical phrase, as all speakers find the sentences in (45) fully grammatical.

(45)a. **Cine**, după părerea ta, a făcut asta?  
*who after opinion your has done that*  
**Who**, in your opinion, did that?

b. După părerea ta, **cine ce** a făcut?  
*In your opinion, who did what?*

c. **Cine ce** a făcut, după părerea ta?  
*Who did what, in your opinion?*

A parenthetical in a nonmultiple question or at the beginning or end of a multiple question is fine, but either within or after the multiple Wh-word group it produces uncertain acceptability at best. In Bulgarian the situation is similar:

(46)a. ?\***Koj**, spored tebe, **kakvo e** казал?  
*who according to you what has said*  
**Who**, in your opinion, said **what**?



- (46)b. ?\***Koj kakvo**, spored tebe, e kazal?  
*who what according to you has said*

This uneasiness with parentheticals in multiple questions does not extend to Serbo-Croatian, Czech, and Polish, however. In Serbo-Croatian, for example, sentences with a parenthetical phrase after the first Wh-word are fine:<sup>14</sup>

- (47)a. **Ko**, po tebi, **šta** pije?  
*who by you what drinks*  
**Who**, according to you, is drinking **what**?
- b. Ono što mene zanima je, **ko**, po tvom mišljenju, **šta**  
*that that me interests is who by your opinion what*  
*radi?*  
*does*  
 What interests me is **who**, in your opinion, is doing **what**?

Cichocki (1983) demonstrates that parentheticals, like clitics, may follow the first Wh-word of a fronted Wh-sequence, or the whole group, but not the second one of three (compare (38) above):

- (48)a. **Kto** wedlug ciebie **komu co** dal?  
*who according to you to whom what gave*  
**Who** in your opinion gave **what to whom**?
- b. \***Kto komu** wedlug ciebie **co** dal?  
*who to whom according to you what gave*
- c. **Kto komu co** wedlug ciebie dal?  
*who to whom what according to you gave*

This could be taken as an indication that the second and third Wh-words form a constituent together (in fact, Cichocki does so argue); however, this conclusion does not hold for the other [-MFS] languages. In Czech, for example, unlike in Polish, parentheticals can appear anywhere in the Wh-word sequence:

- (49)a. **Kdo**, podle tebe, **co komu** dal?  
*who according to you what to whom gave*  
**Who**, according to you, gave **what to whom**?

<sup>14</sup> The parenthetical phrase may come at the end of the Wh-word sequence too, as well as in other positions in the sentence.

- (49)b. **Kdo co, podle tebe, komu dal?**  
*who what according to you to whom gave*

In spite of minor language-specific differences, the distribution of parentheticals, like that of clitics and adverbs, indicates that there is a major constituent break between at least the first and second Wh-words in Serbo-Croatian, Polish, and Czech, while in Bulgarian and Romanian all of the WHs form an indivisible unit. In our analysis the first and only the first WH in the first group of languages has undergone Wh-movement, and it alone is in SpecCP; while in the second group of languages all of the fronted WHs are in SpecCP and therefore form a constituent.

In addition to the evidence from parentheticals and clitics, there are some further arguments for constituent structure in individual multiple-Wh-fronting languages, all of which point to the same conclusion.<sup>15</sup> I will now briefly mention several of these, beginning with an argument from the morphology of definiteness marking in free relatives in Bulgarian. Relative as opposed to interrogative Wh-words in Bulgarian are marked with a definitizing suffix **-to** (compare (50a–b)). In multiple relatives each Wh-word may be separately marked, as in (50b), but this is not necessary. One **-to** at the end of the Wh-word sequence suffices to mark the whole sequence relative (50c). Note that this **-to** must be suffixed to the last Wh-word; as (50d) shows, marking only the first Wh-word is ungrammatical. This strongly suggests that in (50c) **-to** is suffixed to the constituent consisting of the two Wh-words **[[koj kakvo]-to]**:

- (50)a. **Koj kakvo iska?**  
*who what wants*  
**Who wants what?**

<sup>15</sup> In addition, other multiple-fronting languages may provide further types of evidence for constituent structure beyond those that can be seen in the languages studied in this article. For instance, as Ellen Prince has informed me, Yiddish, which appears to be a [+MFS] language, provides a strong indication not available in Slavic that multiple fronted Wh-words are not in IP. A verb-second constraint requires **es** to fill the first position in IP in Wh-questions (for details on pleonastic **es** in Yiddish, see Travis (1984), especially pp. 185–186):

- (i) Ikh veys nit **ver** es iz geshtorbn.  
*I know not who it is died*  
 I don't know **who** died.

In multiple questions **es** is still required and follows all of the WHs:

- (ii) Ikh veys nit **ver vemen** es hot gezen.  
*I know not who whom it has seen*  
 I don't know **who** saw **whom**.

- (50)b. **Kojto kakvoto** iska ...  
*who-DEF what-DEF wants*  
**Whoever wants whatever ...**
- c. **Koj kakvoto** iska ...  
*who what-DEF wants*  
**Whoever wants whatever ...**
- d. \***Kojto kakvo** iska ...  
*who-DEF what wants*

A further argument for the [-MFS] hypothesis comes from Polish sentences like (51) (from Lasnik and Saito), in which a relative and an interrogative Wh-word have been fronted in the same clause. Assuming that the head and specifier, Comp and SpecCP, reflect the [interrogative] status of the clause,<sup>16</sup> both Wh-words cannot be in the same specifier position in this sentence: a clause cannot be both interrogative and noninterrogative (relative). However, if **który** is in SpecCP and **kogo** is neither in SpecCP nor in Comp,<sup>17</sup> but rather in or adjoined to IP, as is the case under the [-MFS] analysis of Polish, there is no problem.

- (51) Spotkałéś mężczyznę, **który kogo** zabił?  
*met-2s the man who whom killed*  
**Whom** did you meet the man **who** killed?

Although this is a somewhat different case from the multiple questions we have been looking at, it does indicate that some fronted Wh-words in Polish are not in SpecCP. Clauses containing a complementizer together with a fronted Wh-word, like the following example from Lasnik and Saito (1984), make the same point.

- (52) Maria powiedziała, **że kto** odwiedza Janka?  
*Maria said that who visits Janek*  
**Who** did Maria say visits Janek?

Not only does the Wh-word here not agree in interrogativity with its clause, it also follows rather than precedes Comp and thus cannot be in SpecCP. The constructions in (51) and (52) are impossible in Bulgarian and Romanian; I do not know if they exist in Czech or Serbo-Croatian.

<sup>16</sup> See Rudin (1986) for some discussion of this assumption.

<sup>17</sup> See below, section 1.4, for further discussion of the fact that the second Wh-word cannot be in Comp.

A third piece of evidence comes from the demonstrative particle **to** in Serbo-Croatian. This particle is not a clitic. It occurs only in main clauses, and is sentence-initial except in Wh-questions, where it follows the fronted Wh-word. Since Wh-questions are the only main clauses with an overt SpecCP in Serbo-Croatian, this indicates that **to** follows SpecCP. Significantly, in multiple questions it follows the first Wh-word (and of course any clitics attached to it), but precedes the second Wh-word:

- (53) **Tko** je **to koga** udario? (Browne 1976)  
*who has that whom hit*

**Who hit whom?/Who was that hitting whom?**

See Browne (1976) for more detailed discussion of this particle and its role in distinguishing two separate Wh-positions in Serbo-Croatian, one for the first Wh-word and another for the second and subsequent Wh-words.

Finally, evidence may be available from intonation. Cichocki (1983) notes that Polish has an intonation break between the first and second Wh-words in a multiple question, but not between the second and third or third and fourth, suggesting once again that the first Wh-word is a constituent separate from the others. Having no data on intonation in the other multiple fronting languages, I do not follow up this line of argument here.

#### 1.4. *Wh-Word Order*

##### 1.4.1. *Fixed and Free Order*

The last area of difference between the two groups of multiple-Wh-fronting languages that I will discuss is the order of fronted Wh-words, which is relatively fixed in the [+MFS] languages but not in the [-MFS] languages. To the extent that it can be attributed to the interplay between the two proposed structures for multiple-Wh-fronting constructions and independent principles such as the ECP, this difference supports the analysis presented. I begin by giving some contrasting data from Bulgarian and Serbo-Croatian. In Bulgarian, a nominative must precede an accusative Wh-word, and when a Wh-word indirect object is also present, the order of the three Wh-words must be subject, direct object, indirect object:

- (54)a. **Koj kogo** vižda?  
*who whom sees*

**Who sees whom?**

(54)b. \***Kogo koj vižda?**

- c. **Koj kogo na kogo e**  
*who whom to whom has*  
 pokazal? (no other word order possible)  
*pointed out*  
**Who pointed out whom to whom?**

In Serbo-Croatian, on the other hand, accusative and nominative Wh-words can occur in either order, as shown in (55), albeit with some difference in emphasis and appropriate context. All possible orders with three Wh-words are grammatical as well; see (56).

(55)a. **Ko koga vidi?**  
*who whom sees*  
**Who sees whom?**

b. **Koga ko vidi?**

- (56)a. **Ko je što kome dao?**  
*who has what to whom given?*
- b. **Ko je kome što dao?**  
 c. **Što je ko kome dao?**  
 d. **Što je kome ko dao?**  
 e. **Kome je što ko dao?**  
 f. **Kome je ko što dao?**  
**Who gave what to whom?**

The conditions on the ordering of various types of Wh-words in Bulgarian are rather complex,<sup>18</sup> but on the whole they match quite neatly the effects of SUPERIORITY in languages like English. These effects have been attributed to properties of Wh-movement at LF in English (see Section 1.4.2). It is, then, a reasonable hypothesis that the difference in word order between the two languages is due to Bulgarian, but not Serbo-Croatian, having the Wh-words in the same structural position and the same relation to their traces at S-structure as English has at LF. Given the claim of this paper that Romanian and Bulgarian are opposed to Serbo-Croatian, Polish, and Czech in precisely this way, it is no

<sup>18</sup> In some cases, for instance, there appear to be differences between human and nonhuman Wh-words with the same grammatical function, and there is some variability across speakers in just how rigid the order of Wh-words is. For further details, see Rudin (1986, chapter 4).

surprise to find that Romanian is similar to Bulgarian, while Polish and Czech follow the Serbo-Croatian pattern with respect to word order. The facts of Romanian, Polish, and Czech are presented very briefly below.

The order of Wh-words in Romanian, as in Bulgarian, is relatively fixed, at least for subjects and objects.<sup>19</sup> Nominative always precedes nonnominative:

- (57)a. **Cine ce** a spus?  
*who what has said*  
**Who said what**
- b. \***Ce cine** a spus?  
*what who has said*
- (58)a. **Cine cu cine** vorbește?  
*who with whom speaks*  
**Who is speaking with whom?**
- b. \***Cu cine cine** vorbește?  
*with whom who speaks*

In Polish as in Serbo-Croatian Wh-word order is quite free, as shown in the following sentence pairs from J. R. Williams (1986b). Note in particular that accusative can precede nominative, in (60b).

- (59)a. **Kogo komu** przedstawiłeś?  
*whom to whom introduced-2s*  
**Whom did you introduce to whom?**
- b. **Komu kogo** przedstawiłeś?  
*to whom whom introduced*
- (60)a. **Kto co** robił?  
*who what did*  
**Who did what?**
- b. **Co kto** robił?  
*what who did*

The order of Wh-words within the fronted Wh-sequence in Czech is

<sup>19</sup> I do not have detailed data on the order of accusative/dative Wh-word combinations in Romanian, but it appears to be less strict than in Bulgarian.

also rather free. Once again, note especially that the nominative need not precede other Wh-words (examples from Toman (p.c.)):

- (61)a. **Kdo kdy koho** pozval, neví  
*who when whom invited I don't know*  
**Who** invited **who when**, I don't know.
- b. **Kdy kdo koho** pozval, neví  
*when who whom invited I don't know*
- c. **Koho kdy kdo** pozval, neví.  
*whom when who invited I don't know*

Thus Wh-word order is free in the [-MFS] languages, but restricted at least to subject > nonsubject in the [+MFS] languages. This word order difference is attributable to the differing structural properties of multiple questions in the two groups of languages, as we shall see below.

#### 1.4.2. *Word Order and Locality Principles*

In the previous section we saw that in Bulgarian and Romanian, if both a subject and an object Wh-word are fronted, the subject must precede the object. This is reminiscent of the superiority<sup>20</sup> effect in English and other non-multiple-fronting languages (e.g. Dutch: Koopman (1984, p. 224)); i.e. of the fact that in multiple questions the subject rather than the object must be moved if both are Wh-phrases:

- (62)a. I wonder **who** did **what**.  
 b. \*I wonder **what who** did.

In recent treatments (Aoun, Hornstein, and Sportiche 1980; Chomsky 1981; and many later works), this subject-object asymmetry is often subsumed under some version of the ECP, the claim being that at LF  $e_i$  is properly governed (under some definition of proper government) in (63a) but not (63b), the LF representations of (62a, b) respectively.

- (63)a. ... [what<sub>j</sub> [who<sub>i</sub>]]  $e_i$  did  $e_j$   
 b. \*... [who<sub>i</sub> [what<sub>j</sub>]]  $e_i$  did  $e_j$

If, as the [+MFS] analysis implies, Bulgarian and Romanian have the same structure for multiple questions at S-structure that other languages have at LF, we might expect the restrictions on Wh-word order in

<sup>20</sup> See Chomsky (1973) for the original Superiority condition.

[+MFS] languages to fall under the ECP or a similar locality principle as well.<sup>21</sup>

Some objections to a classical ECP analysis of these word order facts were presented in Rudin (1985), including, among other problems, the fact that such an analysis would not account for superiority effects beyond the simple subject nonsubject asymmetry (e.g. accusative/dative asymmetries as in (54c)), nor for the relative freedom of occurrence of Wh-phrases containing lexical material other than a simple Wh-word (compare the Bulgarian examples (64a–b)).

- (64)a. Ne razbiraš        **na koja žena kakŭv muž**  
           *NEG understand-2s to which woman what kind man*  
           trjabva.  
           *is necessary*  
           You don't understand **what kind of man which woman** needs.
- b. \*Ne razbiraš        **na kogo koj** trjabva.  
           *NEG understand-2s to whom who is necessary*  
           You don't understand **whom who** needs.

Both problems have been noted for English and other languages as well (Kayne 1980; Fiengo 1980; E. Williams 1986; Comorovski 1986; Contreras 1986; Pesetsky 1987; among others) and have caused some (for instance, AOUN, HORNSTEIN, LIGHTFOOT, AND WEINBERG (1987); henceforth AHLW) to conclude that some, though not all, superiority effects are unrelated to proper government.

On the other hand, though, recent SPLIT versions of the ECP, developed by Travis (1984), Jaeggli (1986), Davis (1987), AHLW, and others can account for at least some of the Slavic facts. Various formulations of the split ECP differ, but all have in common that they break up the old ECP – the requirement that empty categories be properly governed at LF – into two separate conditions – a government require-

---

<sup>21</sup> The distribution of Wh-adjuncts – another phenomenon which has been attributed to the ECP or similar constraints such as Huang's (1982) Condition on Extraction Domains – is also different in the [+MFS] languages than in [-MFS] languages. The facts in this area are subtle and complex, and I leave their untangling for future work. Oversimplifying, the basic facts seem to be that nonsubcategorized adjuncts like *how* and *why* are not permitted in multiple questions in [+MFS] languages at all, except in conjoined structures (*What did you do, and why?*), while in [-MFS] languages they occur somewhat more freely.



ment and a binding requirement – which must both be satisfied, but not necessarily at the same level of the grammar.

For the sake of concreteness, I adopt here the version of the split ECP proposed by AHLW, with minor adjustments to fit an analysis in which Wh-words are in SpecCP rather than Comp. Wh-traces must satisfy two conditions, in this system – head government (65) at PHONOLOGICAL FORM (PF) and a local binding condition (66) (a subcase of Aoun's (1985) Generalized Binding) at LF.

- (65) An empty element must be governed by a lexical head (at PF).  
 (66) An *A'* anaphor must be *A'* bound in its Domain (at LF).

Government is defined as in (67); from AHLW, based on an earlier definition by Aoun and Sportiche:

- (67) *A* ( $X^0$ ) GOVERNS *B* iff all maximal projections dominating *A* also dominate *B* and *A* is dominated either by all maximal projections dominating *B* or by all maximal projections dominating the maximal projection of *B*.

The definition of Domain is very slightly modified from that given in AHLW:

- (68) The DOMAIN for an expression *A* is the first clause (IP or CP) or NP which contains an accessible SUBJECT for *A*, where SUBJECT = AGR, [NP, IP], or [NP, NP] (Chomsky 1981), and where *B* is accessible to *A* iff *A* is in the c-command domain of *B* and assigning the index of *B* to *A* would violate neither the *i*-within-*i* condition nor BINDING CONDITION C (the condition that R-expressions must be *A*-free).

Finally, it should be noted that all and only indexed items are 'visible' at PF; thus (65) applies to indexed empty categories, and an indexed head, including indexed Comp, counts as lexical for (65).

Detailed discussion of how this system applies to the two types of multiple-Wh-fronting languages must be deferred until after the structure of the Wh-word sequence in each type has been determined more exactly, in Sections 2.2 and 2.3. At that point we will see that the unacceptable non-subject > subject Wh-word order in [+MFS] languages is ruled out by (66) and perhaps also (65), while Wh-word order is irrelevant to the satisfaction of these conditions in [-MFS] languages.

## 2. THE TWO S-STRUCTURES FOR MULTIPLE-WH-FRONTING LANGUAGES: TOWARD AN ACCOUNT

### 2.1. *Summary of Facts*

To sum up what we have seen so far, Bulgarian and Romanian, on the one hand, differ systematically from Serbo-Croatian, Polish, and Czech, on the other hand, in several ways. These are summarized in the table below:

(69)		Bulgarian	Romanian	Serbo-Croatian	Polish	Czech
	multiple WH extraction					
	from a clause	+	+	-	-	-
	Wh-island					
	violations	+	+	-	-	-
	clitics follow					
	first Wh-word	-	-	+	+	+
	parentheticals,					
	adverbs, particles					
	after first WH	-	-	+	+	+
	free NOM/ACC					
	Wh-word order	-	-	+	+	+

The clustering of properties in the two groups of languages supports the conclusion that all of these properties are to be attributed to a difference in the setting of a single parameter. We have seen above that this parameter corresponds to the descriptive generalization stated in (70):

- (70) THE MULTIPLY-FILLED SPEC CP HYPOTHESIS:  
SpecCP can contain multiple Wh-words at S-structure in Bulgarian and Romanian ([+MFS] languages), but only one in Serbo-Croatian, Polish, and Czech ([-MFS] languages).

Although both types of languages move all Wh-phrases to the beginning of the clause, the clause-initial landing sites and the resulting S-structures for multiple-Wh-constructions differ. Languages of the [+MFS] type have a structure like (71a) (=2c), while those of the [-MFS] type have one like (71b) (=2d) for multiple questions at S-structure.

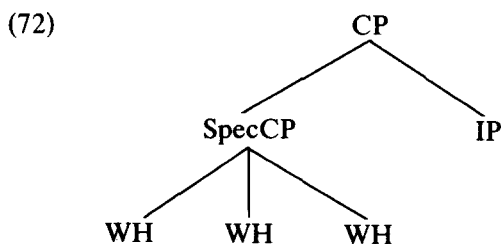
- (71)a.  $[_{CP} [_{SpecCP} WH WH WH] [_{IP} \dots]]$  [+MFS] languages  
b.  $[_{CP} [_{SpecCP} WH] [_{IP} WH WH \dots]]$  [-MFS] languages

But several important questions remain. What is the internal structure of SpecCP in [+MFS] languages? Exactly what position do the Wh-words occupy within IP in the [-MFS] languages? And is it possible to derive (70) from some deeper principle that would allow us to integrate the multiple-Wh-fronting languages into a general typology of Wh-movement? It is to these questions that I now turn. Sections 2.2 and 2.3 treat the structure of the Wh-sequence in [+MFS] and [-MFS] languages, respectively, and Section 2.4 links (70) to a proposal that the level(s) at which adjunction to SpecCP is permitted is subject to language-specific variation, based on the Condition on Comp Adjunction proposed by Adams (1984).

2.2. *The Structure of SpecCP in [+MFS] Languages*

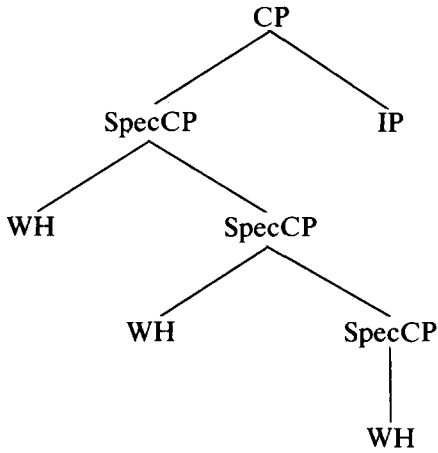
To this point we have seen a number of arguments that all of the Wh-words are in SpecCP at S-structure in Bulgarian and Romanian, but minimal attention has been given to questions relating to the internal structure of multiply-Wh-filled SpecCP. The hierarchical structure of SpecCP, if any, is important in determining how principles of binding and government will apply to [+MFS] languages.

There are several possible structures for a single SpecCP constituent containing multiple Wh-words. One is a flat structure like (72), in which SpecCP dominates an arbitrary number of (presumably base-generated) nodes which can be filled by Wh-movement. This is parallel to the multiply-filled Comp structures assumed for Bulgarian by Rudin (1986) and for Romanian by Comorovski (1986).

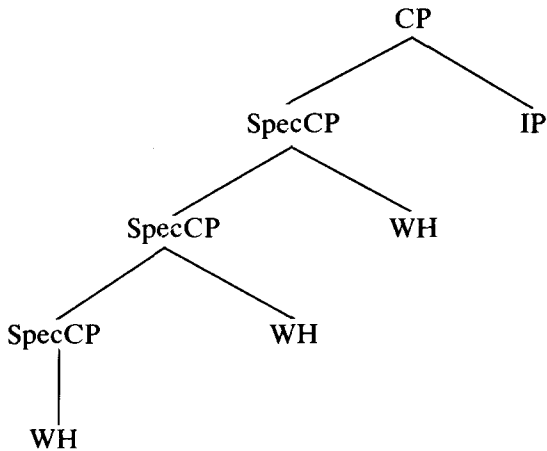


Another possibility is that Wh-words are adjoined to the left or right of SpecCP, as I assumed in the first part of this article, resulting in a structure like (73a) or (73b), respectively.

(73)a.



b.



I argue here for structure (73b). This analysis involves three semi-independent components, which will be discussed in turn:

A. A single Wh-phrase is substituted into the SpecCP position, which is empty at D-structure.

B. An additional Wh-phrase or phrases may be adjoined to SpecCP.

C. The adjunction in (B) is to the right.

An assumption underlying the following discussion is that the structure of multiply-Wh-filled SpecCP at S-structure in the [+MFS] languages is the same as that of multiply-Wh-filled SpecCP at LF universally.<sup>22</sup> That is, I assume that no restructuring of SpecCP takes place between S-structure and LF in [+MFS] languages, although nothing crucial will

<sup>22</sup> Or at least in all languages that permit multiple questions. See section 2.4.

hinge upon this. Following Chomsky (1986), I assume that substitution of a maximal projection into specifier position is one of the possible types of movement, and that syntactic Wh-movement in English, as well as in non-multiple questions in [+MFS] languages, as in (74), is an instance of this type.

- (74) [<sub>SpecCP</sub> kakvo<sub>i</sub>] pravi Ivan t<sub>i</sub> (Bulgarian)  
           *what does Ivan*  
           What is Ivan doing?

This follows directly from the adoption of the system of phrase structure in Chomsky (1986). It is only a very small step from this to the further assumption that one of the Wh-phrases in a multiple question is substituted into SpecCP in Bulgarian and Romanian as in English, i.e. (A) above.

Component (B) of the analysis is somewhat less obvious. Chomsky (1986) does not state what position multiple Wh-phrases would occupy at LF in his system; in fact, he explicitly declines to detail “the specific mechanisms of LF movement of Wh-in-situ” (p. 49). But nothing in the Barriers model prevents adjoining Wh-phrases to SpecCP: SpecCP is XP and is not an argument position, so it should be a permissible adjunction site. I take an adjoined structure to be preferable to a flat structure like (72) on theoretical grounds, since it allows us to restrict movement to two highly constrained types: substitution into an empty node, and adjunction, which Chomsky limits to nonargument XP. Tree (72) requires substitution into a node that already contains one or more Wh-phrases. Furthermore, an adjoined structure permits a simple account of subject/object asymmetries based on the split ECP (see above), as I will now demonstrate.<sup>23</sup>

Recall that a subject Wh-phrase must precede a nonsubject in multiple questions in Bulgarian and Romanian:

- (75)a. **Koj kakvo** pravi? (Bulgarian)  
           *who what does*  
           **Who** is doing **what**?

<sup>23</sup> AHLW argue against any hierarchical structure in Comp (where they place Wh-words) on the grounds that Wh-words need to c-command their traces to satisfy quantification theory. But cf. Chomsky’s (1986) claim that Wh-words in multiple questions need not and in fact do not govern their traces at LF.

- (75)b. \***Kakvo koj** pravi?  
*what who does*  
 \***What is who** doing?

I propose that the single WH which is substituted into SpecCP becomes the head of SpecCP, and its index percolates to SpecCP. By specifier/head agreement (Chomsky 1986), SpecCP passes its index to Comp, the head of CP. This has the effect of coindexing Comp with a fronted Wh-phrase, in a fashion analogous to the Comp-indexing mechanism proposed by Aoun, Horstein, and Sportiche (1980) and adopted in revised form by Lasnik and Saito (1984) and others.<sup>24</sup> Question (75a) has the indexed structure (76), assuming that Wh-words are adjoined to the right of SpecCP:

- (76) [<sub>SpecCP<sub>i</sub></sub> [<sub>SpecCP</sub> koj<sub>i</sub>] kakvo<sub>j</sub>] [Comp<sub>j</sub>] [<sub>IP</sub> t<sub>i</sub> ... t<sub>j</sub> ...]

The indexed Comp, an X<sup>0</sup> category, counts as a visible head at PF in the AHLW system of proper government I am assuming here (outlined above; section 1.4.2). Comp thus licenses t<sub>i</sub> by head-governing it. The object trace of course is head-governed by V. In (75b), with structure (77), Comp is coindexed with the object, not the subject Wh-phrase, so the subject trace appears not to be head-governed and hence the sentence is ruled out:

- (77) [<sub>SpecCP<sub>j</sub></sub> [<sub>SpecCP</sub> kakvo<sub>j</sub>] koj<sub>i</sub>] [Comp<sub>j</sub>] [<sub>IP</sub> t<sub>i</sub> ... t<sub>j</sub> ...]

However, there is an alternative: t<sub>i</sub> in (77) might actually be head-governed by INFL. INFL is a head, and is coindexed with the subject of its clause. If, as has been claimed for Spanish (Chomsky 1981), free inversion correlates with proper government by INFL, it might be expected that in Bulgarian and Romanian, which do have Romance-type free inversion, INFL should be able to properly govern (i.e. head-govern) the subject position. The word order in (75b) would then have to be ruled out in some other way. As we shall see in a moment, it is ruled out at LF in any case.<sup>25</sup> In the discussion of the [-MFS] languages it will become clear that there are reasons to prefer the analysis in which INFL does head-govern a subject trace.

Let us turn now to the local binding requirement for anaphors at LF.

<sup>24</sup> A similar proposal for extending Aoun, Horstein, and Sportiche's Compindexing algorithm to a SpecCP analysis has been put forward by Jaeggli (1986).

<sup>25</sup> Similarly, in languages like English, which have no *ec* in subject position in the (ungrammatical) gloss of (75b), the subject/object asymmetry will be accounted for by the LF binding condition alone, regardless of head-government.

In (76), which is the LF as well as the PF structure for (75a),  $t_i$  is A'-bound by the indexed Comp. The object trace,  $t_j$ , has no accessible SUBJECT, hence no Domain, since coindexing it with either AGR or the subject NP would violate Condition C of the Binding Theory. It therefore need not be bound, and the binding requirement is satisfied. In (75b), on the other hand, the binding requirement is not met. The subject trace,  $t_i$  in (77), has AGR as its accessible SUBJECT, and must therefore be A'-bound in CP, its Domain. But it is not bound either by Comp, which is not coindexed with it, nor by the Wh-word *koj*, which does not c-command it. Thus both parts of the split ECP – head government at PF and local binding at LF – are satisfied by (75a), but (75b) fails to satisfy at least the LF condition, given a structure in which the leftmost Wh-phrase is the head of SpecCP, as is assured by the adjunction analysis I am proposing.

The reasons for postulating adjunction of Wh-phrases to the right rather than the left of SpecCP should now be clear. If adjunction proceeded to the left, the word order facts should be exactly the opposite of what they are: the subject Wh-phrase should be restricted to the rightmost, rather than the leftmost, position in the Wh-word group, as the S-, PF, and LF structure (78), not (76), would license a subject Wh-trace:

(78)  $[_{\text{SpecCP}} \text{ kakvo}_j [_{\text{SpecCP}} \text{ koj}_i]] [_{\text{Comp}}] [_{\text{IP}} t_i \dots t_j \dots]$

Given the assumption that syntactic and LF movement to SpecCP are parallel, this suggests that LF adjunction of Wh-phrases to SpecCP in languages like English is to the right too, i.e. that (79b) rather than (79c) is the correct LF structure for (79a), although this difference in directionality does not crucially affect any analysis of English that I am aware of.

- (79)a. Who said what?
- b. [[who] what] said
- c. [what [who]] said

Aoun, Hornstein, and Sportiche, in their original discussion of adjunction to Comp (1980), assumed adjunction to the right, but later analyses have generally assumed adjunction to the left.

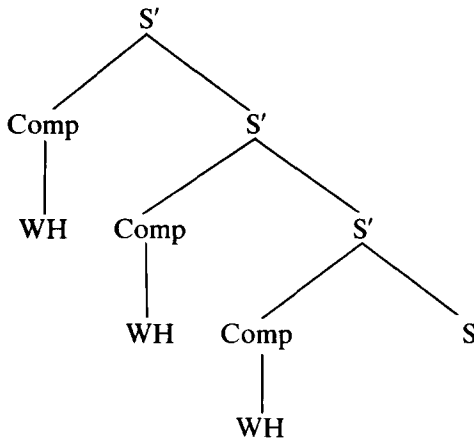
### 2.3. *Adjunction to IP in [-MFS] Languages*

I turn now to the details of Wh-position in the [-MFS] languages – Serbo-Croatian, Polish, and Czech. In the first part of this article I

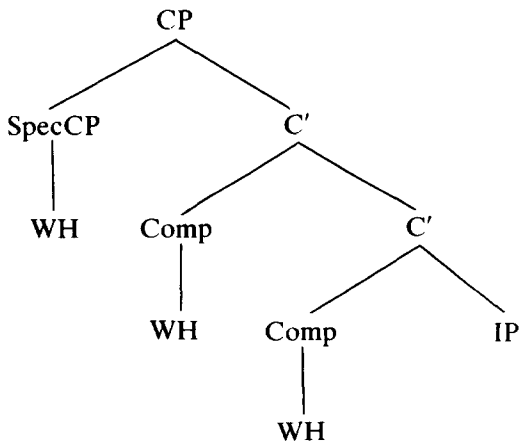
argued that multiple questions in these languages have only one Wh-phrase in SpecCP at S-structure. This Wh-phrase is presumably substituted into the empty specifier position in both single and multiple questions, just as it is in English, the [+MFS] languages, and all other languages with syntactic Wh-movement. The remaining Wh-phrase or -phrases, however, although fronted, are not in SpecCP, but in some other position.

Several possibilities exist for what this position might be; in fact, at least three structures compatible with the general analysis just outlined have been put forward in the literature. Toman (1981) proposed a structure with each Wh-word in a separate Comp, as in (80a), for Czech and Polish; in the X-bar system employed here this might translate to something like (80b):

(80)a.



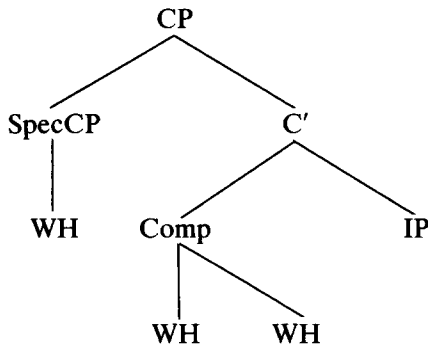
b.



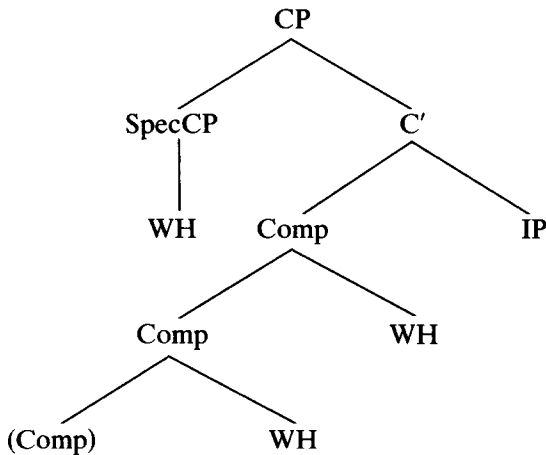


Cichocki (1983), citing evidence from parenthetical and clitic placement (see section 1.3 above) that the second and third Wh-words in a Polish triple question may form a constituent, suggests that the Wh-words other than the first one are in a second Comp (Comp<sub>2</sub>) which differs from the main Comp (Comp<sub>1</sub>) in not being an extraction position and in accommodating multiple Wh-words. To adapt this proposal to the current X-bar framework we might identify Comp<sub>1</sub> with SpecCP and Comp<sub>2</sub> with Comp, expanded in either of the ways shown in (81).<sup>26</sup>

(81)a.



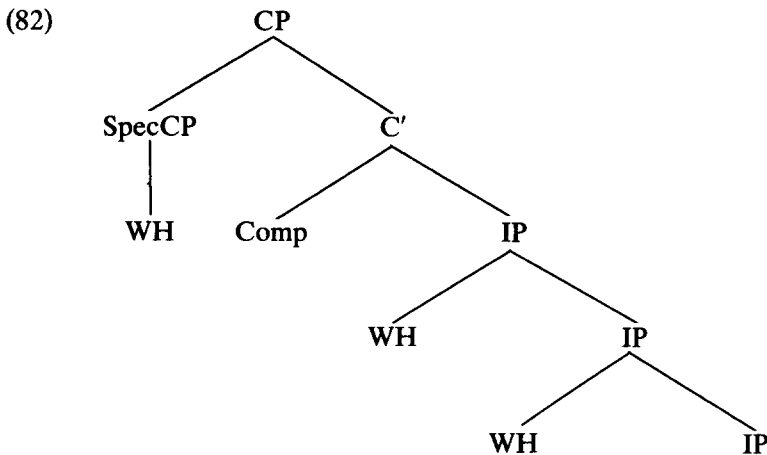
b.



Finally, Lasnik and Saito (1984) suggest a structure for Polish in which each of the noninitial Wh-words is adjoined to S. In the current framework this would be adjunction to IP, producing a tree like (82). It is this

<sup>26</sup> The adjoined structure in (81b) is designed to accommodate sentences like (52), in which a complementizer co-occurs with and precedes a fronted Wh-word; this is why I have shown WHs adjoined to the right, not the left, of Comp. In (80) as well one of the Comps could contain a complementizer. Such sentences are discussed further below.

structure that I will argue is the correct one.



The main difference between (82) and the other structures is that in (80)–(81), Wh-phrases other than the initial one are in Comp, while in (82) they are not. This distinction is vital in light of the pattern of co-occurrence of Wh-words with complementizers in the [–MFS] languages. We have seen that a complementizer can co-occur with and precede a Wh-word:

- (83) Maria myśli, **że co** Janek kupił? (Polish)  
*Maria thinks that what Janek bought*  
**What** does Maria think **that** Janek bought?

However, no other combination of WHs and complementizers is possible; that is, we do not find WH+ complementizer (e.g. in Polish \***co że** ‘what that’), nor, in multiple questions, WH + complementizer + WH (e.g. \***kto że co** ‘who that what’). The facts of the other [–MFS] languages are analogous.<sup>27</sup> The possible and impossible combinations and their structures according to hypothesis (82) are listed in (84):

- (84)
- |             |  |
|-------------|--|
| WH WH       | [ <sub>SpecCP</sub> WH][ <sub>Comp</sub> e][ <sub>IP</sub> WH [ <sub>IP</sub> . . . ]]     |
| Comp WH     | [ <sub>SpecCP</sub> e][ <sub>Comp</sub> Comp][ <sub>IP</sub> WH [ <sub>IP</sub> . . . ]]   |
| *WH Comp    | *[ <sub>SpecCP</sub> WH][ <sub>Comp</sub> Comp][ <sub>IP</sub> [ <sub>IP</sub> . . . ]]    |
| *WH Comp WH | *[ <sub>SpecCP</sub> WH][ <sub>Comp</sub> Comp][ <sub>IP</sub> WH [ <sub>IP</sub> . . . ]] |

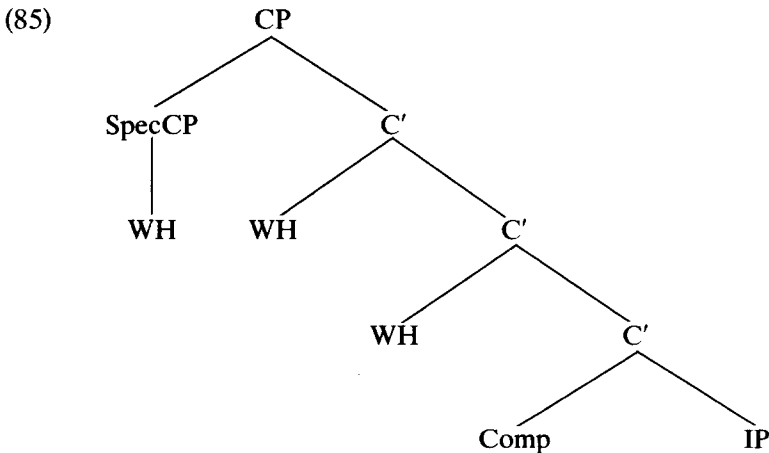
A clear generalization emerges from this array. Apparently the [–MFS] languages, like modern English, obey a form of the DOUBLY-

<sup>27</sup> With the possible exception of one construction in Serbo-Croatian in which Wh-words are treated as indefinite pronouns; this will be discussed below.

FILLED COMP FILTER: either SpecCP or Comp must be empty. This simple account of the distribution of complementizers depends on accepting (82) as the structure of multiple questions in the [-MFS] languages, since under either (80) or (81), the generalization would not hold, SpecCP and Comp both being filled in multiple questions.

Another argument against the second WH being in Comp, based on the co-occurrence of relative and interrogative Wh-phrases, has already been sketched in connection with example (51) above.

Furthermore, the structures in (80)–(81) present some additional problems. The primary rationale for (81) – the claim that the second and third WHs form a constituent – receives no support in Czech or Serbo-Croatian, as we have already seen (section 1.3). Structure (81a) is undesirable for theoretical reasons as well: the two WHs dominated by Comp could not be placed here by either of the two standard types of movement – substitution or adjunction.<sup>28</sup> And (80b) is similarly awkward: it would require an adjoined structure with an appropriate number of Comp positions to be present at D-structure, so that Wh-phrases could be substituted into them. While base generation of adjoined structures is not impossible, they are normally created by movement. Adjunction of WH directly to C', as in (85), would sidestep this drawback, but at the expense of creating further problems, both empirical and theory-dependent. First, it is adjunction to a nonmaximal category, ruled out by Chomsky (1986). And second, it predicts that any complementizer should follow all fronted Wh-words, which we have just seen is false.



<sup>28</sup> See above, and Chomsky (1986). Also, for arguments that branching must be binary, see e.g. Kayne (1984).

Structure (82), with Wh-words adjoined to IP, avoids all of these problems, while still placing the noninitial Wh-words in multiple questions in an A' position other than SpecCP. The fact that each WH A'-binds its trace is crucial for our account of the lack of subject-object superiority effects in Serbo-Croatian, Polish, and Czech. Recall that in [-MFS] languages, unlike [+MFS] languages, Wh-word order is free. Both (86a) and (86b) are grammatical, so each must satisfy both parts of the split ECP.

- (86)a. **Ko koga** vidi? (Serbo-Croatian)  
*who whom sees*
- b. **Koga ko** vidi?  
*whom who sees*

The PF (and S-) structure of these sentences is shown in (87a–b). PF is the level at which the head-government condition must be met.

- (87)a. [<sub>SpecCP<sub>i</sub></sub> ko<sub>i</sub>][Comp<sub>i</sub>]<sub>[IP koga<sub>j</sub> [IP t<sub>i</sub> ... t<sub>j</sub> ... ]]</sub>  
 b. [<sub>SpecCP<sub>j</sub></sub> koga<sub>j</sub>][Comp<sub>j</sub>]<sub>[IP ko<sub>i</sub> [IP t<sub>i</sub> ... t<sub>j</sub> ... ]]</sub>

In (87a),  $t_i$  is head-governed by the coindexed Comp, and  $t_j$  by V, so (86a) is acceptable at PF. Structure (87b) is slightly more problematical. Here Comp head-governs  $t_j$ , but it is not immediately obvious what governs  $t_i$ , the subject trace. There are two alternatives: either  $t_i$  is head-governed by INFL, as discussed for the [+MFS] languages above, or else the definition of head-government could be modified in such a way that the Wh-word **ko** adjoined to IP could serve as a head-governor of its trace. The latter option seems undesirably ad hoc, since **ko**, although lexical, is not in any normal sense a head. I therefore tentatively adopt the position that INFL, which is a head and which counts as a lexical category because of its coindexation with the subject NP, head-governs a subject trace at PF in both the [-MFS] and the [+MFS] languages discussed in this article.

At LF, the Wh-phrase(s) adjoined to IP move to SpecCP. So (88a–b) are the LF structures of (86a–b), respectively:

- (88)a. [<sub>SpecCP<sub>i</sub></sub> [ko<sub>i</sub>] koga<sub>j</sub>][Comp<sub>i</sub>]<sub>[IP t<sub>j</sub> [IP t<sub>i</sub> ... t<sub>j</sub> ... ]]</sub>  
 b. [<sub>SpecCP<sub>j</sub></sub> [koga<sub>j</sub>] ko<sub>i</sub>][Comp<sub>j</sub>]<sub>[IP t<sub>i</sub> [IP t<sub>i</sub> ... t<sub>j</sub> ... ]]</sub>

Object Wh-traces are not subject to the local binding requirement at LF, since they have no Domain, coindexation with either AGR or [NP, IP] being a Condition C violation. The trace  $t_i$  in subject position in both (88a) and (88b) has CP as its Domain, and in both cases it is A'-bound

within that Domain: by the indexed Comp (and/or SpecCP) in (88a) and by  $t_i$  adjoined to IP in (88b). The trace adjoined to IP is not A'-bound in either (88a) or (88b). However, like the object trace, it has no accessible SUBJECT, hence no Domain, hence does not fall under the local binding requirement.<sup>29</sup> So both (88a) and (88b) are legitimate LF structures; unlike in the [+MFS] languages neither word order in (86) is ruled out at LF.

Finally, a possible objection to structure (82) should be noted. Chomsky (1986) disallows adjunction of Wh-phrases to IP, for reasons internal to his Barriers treatment of subadjacency and related phenomena. (Other elements, such as quantifiers, do adjoin to IP.) This need not be of great concern here, since I am not assuming the Barriers account of these matters. Nonetheless, it is worth considering one last alternative which would avoid adjunction to IP, namely, adjunction to VP. A question like (89a) would have the structure shown in (89b) rather than (89c), where  $e$  is the null subject of IP.

(89)a. **Kogo komu** przedstawiłeś? (Polish)

*whom to whom introduced-2s*

**Whom** did you introduce **to whom**?

b. [<sub>SpecCP</sub> kogo][<sub>IP</sub> e [<sub>VP</sub> komu [<sub>VP</sub> przedstawiłeś]]]

c. [<sub>SpecCP</sub> kogo][<sub>IP</sub> komu [<sub>IP</sub> e [<sub>VP</sub> przedstawiłeś]]]

But this possibility is straightforwardly ruled out by multiple questions with an overt subject. The following Polish example is typical for all the [-MFS] languages: whether the subject is a pronoun or a nonpronominal NP, it cannot precede the second Wh-phrase.

(90)a. **Kogo komu** Jan/on przedstawił?

*whom to whom Jan he introduced*

**Whom** did Jan/he introduce **to whom**?

b. \***Kogo** Jan/on **komu** przedstawił?

*whom Jan he to whom introduced*

I take it as established then that (82) is the correct structure for multiple questions in the [-MFS] languages.

<sup>29</sup> In Lasnik and Saito's (1984, pp. 280–81) account of Polish, Wh-traces adjoined to S had to delete at LF, not being properly governed. In the present account this is not necessary precisely because they have no Domain.

#### 2.4. *Adjunction to SpecCP at S-structure, LF, and PF*

We have now seen that the differences between the two groups of multiple-Wh-fronting languages reduce essentially to the fact that Wh-phrases adjoin to SpecCP at S-structure in [+MFS] but not [-MFS] languages (where they adjoin instead to IP). Given the universality of Wh-raising at LF (Aoun, Horstein, and Sportiche 1980; Huang 1982; Chomsky 1986, p. 50), this in turn reduces to a difference in the level at which adjunction to SpecCP can apply. In both types of languages WHs adjoin to SpecCP; the only difference is that in [+MFS] languages like Bulgarian adjunction takes place at or before S-structure, while in [-MFS] languages like Serbo-Croatian it occurs at LF. A proposal to account for certain properties of Wh-movement and multiple interrogation in several non-multiple-Wh-fronting languages in terms of language-specific differences in the levels at which adjunction can occur has been put forward by Adams (1984). In this section I demonstrate that her proposal, with minor modifications, can be extended to cover the multiple-Wh-fronting languages as well.

Adams' suggestion consists of a prohibition on adjunction to Comp at specific levels of the grammar; languages differ in what, if any, levels (91) applies at.

- (91)    CONDITION ON COMP ADJUNCTION (CCA) (Adams 1984)  
           \*<sub>[Comp α Comp]</sub>  
           (i.e. nothing may be adjoined to Comp (at level *X* of the grammar))

To conform to the decision to treat multiple Wh-movement as adjunction to SpecCP, let us adopt instead a condition of SpecCP adjunction:

- (92)    CONDITION ON SPEC CP ADJUNCTION (CSA)  
           \*<sub>[SpecCP α SpecCP]</sub>  
           (nothing may be adjoined to SpecCP)

Adopting (92) rather than (91), along with the assumption that Wh-words are in SpecCP, not Comp, does of course induce some changes in what will count as violations of the condition.<sup>30</sup> However, all of Adams' important insights are preserved. These insights have to do with the relation between multiple questions and Wh-island violations, both of

<sup>30</sup> For instance, Middle English relatives of the type NP [WH that S], which Adams, assuming WH adjoined to Comp, cites as a PF violation of (91), do not violate (92).

which involve Wh-adjunction to SpecCP at some level.<sup>31</sup> Adams proposes that, depending on the level(s) at which CCA applies, a language may have both multiple questions and Wh-island violations, or only one or the other, or neither. Italian, as described by Rizzi (1982), has no multiple questions: the CCA holds at LF. However, Italian does permit certain kinds of Wh-island violations, indicating that the configuration  $WH_i \dots [_{Comp} t_i [_{Comp} WH_j]]$  is permitted at S-structure: CCA does not hold at S-structure. (Adams suggests that the trace  $t_i$  in this structure is invisible to the CCA at LF iff it is a relative operator, thus accounting for the fact that only questioning, not relativization, produces Wh-island violations. We will see below that this suggestion is problematical.) English has the opposite distribution. Multiple questions (with WH in situ) do occur, so CCA does not hold at LF, but Wh-island violations do not occur, so CCA does hold at S-structure. In modern Irish the CCA holds at both LF and S-structure: Irish allows neither multiple questions nor Wh-island violations (McCloskey 1979). As a language in which CCA does not apply at either S-structure or LF, Adams cites the variety of Spanish described by Torrego (1984), which permits both multiple questions and Wh-island violations. Spanish does obey the CCA at PF, however, assuming that Wh-trace is not visible at PF, since overt multiply-filled Comps do not occur.

This analysis of the distribution of multiple questions and Wh-islands in terms of the CCA applying at LF, at S-structure, at both, or only at PF works equally well under the revised Condition on SpecCP Adjunction. In fact, the CSA avoids one problem of the CCA, namely, the role of complementizers in the system. The CCA predicts, apparently falsely, that languages like Middle English, in which a Wh-word and a complementizer may co-occur, should permit multiple Wh-adjunction to Comp at PF: (93a) and (93b) have the same structure at PF even if one assumes the neutral complementizer is not present at earlier levels.

- (93)a.  $[_{Comp} WH [_{Comp} that]]$   
 b.  $[_{Comp} WH [_{Comp} WH]]$

The SpecCP analysis of Wh-movement, though, makes no such prediction; the presence of a complementizer is irrelevant to the structure of SpecCP, and the fact that (94a) occurs does not imply (94b) is possible at that level.

<sup>31</sup> For a much more detailed treatment of the facts and issues discussed in the remainder of this paragraph, see Adams (1984).

- (94)a. [<sub>SpecCP</sub> WH][<sub>Comp</sub> that]  
 b. [<sub>SpecCP</sub> WH [<sub>SpecCP</sub> WH]] ([<sub>Comp</sub>])

Adams briefly mentions the existence of multiple Wh-fronting in Polish and Romanian and suggests tentatively that such languages might represent the option of CCA not applying at any level of the grammar. As I have demonstrated in detail above, these two languages behave quite differently with respect to Wh-adjunction to SpecCP. Adams' conjecture turns out to be correct for Romanian (and Bulgarian) but not for Polish and the other [-MFS] languages.

The [+MFS] languages have multiple Wh-questions, which indicates that CSA does not hold at LF. They also allow Wh-island violations, as we saw in section 1.2, so CSA does not hold at S-structure either. The possibility of multiple Wh-extraction from a clause (section 1.1) is further evidence that CSA is not in force at S-structure, since multiple extraction produces S-structures with Wh-trace as well as Wh-phrases adjoined to SpecCP:

- (95) [[<sub>SpecCP</sub> [<sub>SpecCP</sub> WH<sub>i</sub>] WH<sub>j</sub>] ... [<sub>CP</sub> [<sub>SpecCP</sub> [<sub>SpecCP</sub> t<sub>i</sub>] t<sub>j</sub>] ... t<sub>i</sub> t<sub>j</sub> ... ]]

Furthermore, CSA does not hold at PF in Bulgarian and Romanian. There is no reason to believe that fronted multiple Wh-words in these languages are not adjoined to SpecCP at PF, just as they are at S-structure. They can even co-occur with a neutral complementizer (cf. (94b) above), in emphatic questions like (96):

- (96) Kazaha, **koj kakvo** će e vidjal? (Bulgarian)  
*said-3p who what that has seen*  
 They said that **who** saw **what**?

Given the rather marginal and echo-like character of this construction, I hesitate to draw strong conclusions from it; nonetheless it seems clear that the [+MFS] languages do not obey the CSA, at any level.

The [-MFS] languages also do not obey CSA at LF, since they permit multiple questions. However, unlike the [+MFS] languages, they do observe it at S-structure, since Wh-islands are not violated. (And we have seen several other types of evidence that only one Wh-word is in SpecCP at S-structure as well.) It is less clear whether the CSA holds at PF in these languages. In Serbo-Croatian there is one construction in which two fronted Wh-words may precede a complementizer, suggesting that adjunction to SpecCP does occur at PF.



- (97) Nemamo **što kome** da damo.  
*don't have-1p what to whom that give-1p*  
 We don't have **anything** to give to **anyone**.

However, this is unlike other Wh-constructions in some ways, including the interpretation of the Wh-words as indefinite, not interrogative, pronouns, so it is unclear how much weight to give this evidence. I do not know of evidence for adjunction to SpecCP at PF in Polish or Czech.<sup>32</sup> If varieties of Czech actually exist that allow true Wh-island violations (see section 1.3), they provide theory-internal reason to believe adjunction is disallowed at PF at least in these dialects. These varieties of Czech would presumably be like Spanish, in that CSA does not hold at S-structure, but multiple lexical WHs in SpecCP at S-structure (as opposed to one Wh-phrase and one trace) are ruled out by the CSA applying at PF.

One element of Adams' proposal which turns out to be untenable is the idea that the CCA (or CSA) distinguishes interrogative and relative operators at LF, to account for the fact that relative but not interrogative Wh-phrases can be extracted from Wh-islands in Italian. As we have seen (section 1.2), Bulgarian and Romanian exhibit a similar relative-interrogative asymmetry, although the CSA does not apply at all in these languages. And the same type of asymmetry has been noted for Scandinavian languages (Maling 1978), Hungarian (Horvath 1981), Polish, where questioning gives more severe island violations than relativization (Cichocki 1983), and even English, where many speakers find (98a) better than (98b), although none of these languages obey the CSA at LF.

- (98)a. ?Bill was the one who we all wondered who would ever marry.  
 b. \*Who did you wonder who would ever marry?

I therefore assume that all empty operators (not just relative operators) are invisible to the CSA at LF and that the very common relative/interrogative asymmetry is due to independent factors.

The levels at which the CSA holds in various languages are summarized in the table below. Parentheses around PF markings indicate that this has not been adequately studied. As noted above, the [-MFS]

<sup>32</sup> If a subject Wh-trace is head-governed by INFL in these languages (section 2.3), it makes no difference to the head-government condition whether the Wh-word(s) adjoined to IP move to SpecCP at PF or not.

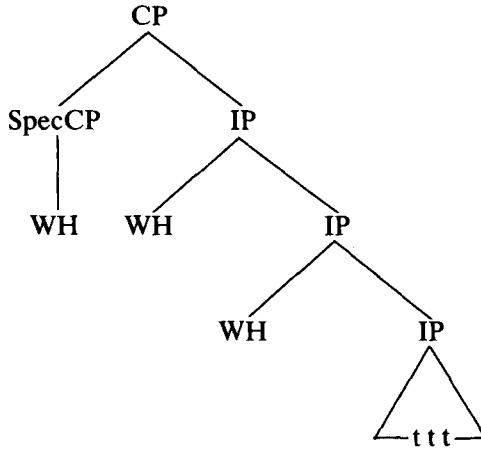
languages may differ from others of type 3 in not respecting the CSA at PF, but this is far from clear.

(99)	CSA applies at:	PF	S-structure	LF
1. [+MFS] languages:				
	Bulgarian	-	-	-
	Romanian (Yiddish)			
2. Spanish (varieties of Czech??)				
		+	-	-
3. [-MFS] languages:				
	Serbo-Croatian			
	Polish	(-)		
	Czech		+	-
	English			
	French	(+)		
	Chinese			
4. Italian				
		(+)	-	+
5. Modern Irish				
		(+)	+	+

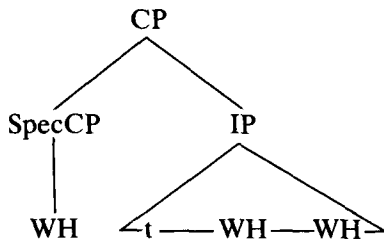
### 2.5. *Some Further Issues*

The grouping of languages in table (99) does not correspond in as direct a way as one might suppose to the S-structure position of Wh-words in multiple questions, the property with which we began this study. Although the [+MFS] languages all have the same CSA settings, languages with one Wh-phrase in SpecCP at S-structure are distributed over four separate CSA configurations. In addition, the languages under 3 – those which obey CSA at S-structure but not at LF – are very diverse, having at least the range of variation in S-structure Wh-positions shown in (100):

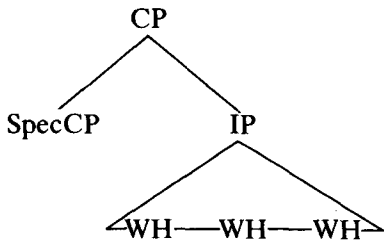
(100)a. Serbo Croation,  
Polish, Czech



b. English<sup>33</sup>  
(French)



c. Chinese  
(French)



Obviously, some further factors are involved in accounting for the existing range of structures for multiple Wh-constructions and related matters. I merely suggest these factors here, leaving a full account for

<sup>33</sup> This of course is also the S-structure for Spanish, which, although it permits Wh-island violations and has multiple questions, adjoins all Wh-words in multiple questions to SpecCP only at LF, because of the ban on multiple Wh-phrases in SpecCP at PF.

further research. The questions raised by (99) and (100) have to do with optionality and obligatoriness of various kinds of Wh-movement: why is syntactic Wh-movement (substitution) to SpecCP optional in French, but obligatory in other languages, and disallowed in the Chinese type? Is movement to non-SpecCP A' positions, as in (100a), obligatory, and if so, why? Why do Wh-words in the [-MFS] languages (as well as [+MFS]) need to be in A' positions at S-structure? Finally, why is adjunction to SpecCP required at S-structure in the [+MFS] languages?

One factor which will probably have to be stipulated for each language is the level(s) at which Wh-movement applies. A given language may have only LF Wh-movement (the Chinese type), only syntactic Wh-movement (Bulgarian and Romanian), or both. The lack of LF Wh-movement in the [+MFS] languages means that WH in situ does not occur: in both Romanian and Bulgarian all Wh-words must be fronted at S-structure to make a nonecho multiple question. (\*marks echo status):

(101)a. **Cine ce** a spus. (Romanian)

*who what has said*

**Who said what?**

b. \***Cine** a spus **ce?**

*Who said what?*

c. \***Ce** a spus **cine?**

*What did who say?*

Furthermore, the CSA cannot apply at S-structure, since Romanian and Bulgarian do have multiple questions. A language with no LF Wh-movement and not allowing adjunction to SpecCP at S-structure would have no way to achieve adjunction to SpecCP at LF, and thus could not have any multiple-Wh-constructions.

Another factor involved may be [+WH] marking mechanisms. It has been suggested (Aoun, Hornstein, and Sportiche 1980; see also Huang 1982 and Aoun 1986) that Comp of an interrogative clause must be marked [+WH] at S-structure for subcategorization checking in languages with syntactic Wh-movement. The optionality of Wh-movement in French can be accounted for if, in French, matrix Comp can be intrinsically marked [+WH],<sup>34</sup> while in English Comp is [+WH] only if it is coindexed with a Wh-phrase. As we have seen, the index of a

<sup>34</sup> But only matrix Comp; syntactic Wh-movement is obligatory in embedded questions. Cf. Aoun (1986, p. 26).

Wh-phrase in SpecCP percolates to SpecCP and then, by Spec/Head agreement, to Comp; Comp must be in configuration (102) to be [+WH] in English.

- (102) [<sub>SpecCP<sub>i</sub></sub> WH<sub>i</sub>] Comp<sub>i</sub>

Thus in English WH in situ is possible only if SpecCP contains a Wh-phrase at S-structure. But in French, since Comp can be intrinsically [+WH], all of the WHs in a multiple question can be in situ, as in (1c).

Intrinsic [+WH] marking is apparently available in Polish too, since movement at LF can be to an empty SpecCP. Recall that Wh-phrases cannot be extracted from finite clauses in the syntax. Co ‘what’ in (103) is adjoined to the lower IP at S-structure, but moves to SpecCP of the matrix clause at LF.<sup>35</sup> Thus the difference in [+WH] marking cuts across the multiple fronting/non-multiple fronting distinction.

- (103)a. *Maria myśli, że co Janek kupił?*  
*Maria thinks that what Janek bought*  
**What** does Maria think Janek bought?

- b. LF: [<sub>SpecCP</sub> CO<sub>i</sub>] Maria myśli [<sub>CP</sub> że [<sub>IP</sub> t<sub>i</sub> [<sub>IP</sub> Janek kupił]]]

It remains to discuss the issue of how to account for IP-adjunction, and whether or not it is obligatory in the [-MFS] languages. The answer to the latter question is that the [-MFS] languages differ among themselves. In Polish and Czech, WH in situ exists only with echo interpretation.<sup>36</sup> In (104)–(105), only the (a) versions, with all WHs fronted, are true multiple questions.

- (104)a. **Kto kiedy wyjechał?** (Polish)  
*who when left*  
**Who left when?**

- b. \***Kto wyjechał kiedy?**  
*who left when*

<sup>35</sup> Lasnik and Saito (1984) saw sentences like (103) as a problem, since they seemed to require Comp-to-Comp movement at LF, which otherwise did not occur. This problem disappears in the present analysis, since *co* at S-structure is not in Comp (or SpecCP).

<sup>36</sup> Also perhaps with what Pesetsky (1987) calls a “D(iscourse)-linked” reading; he suggests that WH in situ in Polish does not undergo LF movement.

- (105)a. **Kdo koho** viděl? (Czech)  
*who whom saw*  
**Who saw whom?**

- b. \***Kdo** viděl **koho**?  
*who saw whom*

In Serbo-Croatian, however, nonecho interpretation of WH in situ in (106b) is possible for many speakers.

- (106)a. **Ko** je **koga** video?  
*who has whom seen*  
**Who saw whom?**

- b. **Ko** je video **koga**?  
*who has seen whom*

Differences in acceptability of Wh in situ may be linked to constraints on LF Wh-movement. Polish apparently allows LF movement only from A positions, and Serbo-Croatian from both A' and A positions. Whether this pattern of LF movement can be derived from more general principles is unclear, however.

### 3. CONCLUDING SUMMARY

My goals in this article have been to describe two types of multiple-Wh-fronting languages and to integrate them into a general typology of Wh-movement and multiple-Wh-constructions. In the first part of the article I demonstrated that in spite of their superficial similarity, the East European languages which normally place all Wh-words at the beginning of the clause fall into two distinct groups with different structures for multiple-Wh-constructions. One group, which includes Bulgarian and Romanian, places all of the Wh-words of a multiple question in SpecCP at S-structure, while the second group, including Polish, Czech, and Serbo-Croatian, has only one Wh-word in SpecCP. I have referred to these two groups as the [+MFS] (Multiply Filled Specifier) and [-MFS] languages, respectively. The difference in S-structure accounts for systematic differences in extraction possibilities for multiple Wh-words, island effects, clitic position and occurrence of parentheticals and other material within the fronted Wh-word sequence, and strictness of word order in multiple-Wh-constructions in the two types of languages.

The second part of the article attempts to locate the two types of multiple-Wh-fronting languages within the broader range of possible

structures for multiple questions and to provide some account of this range. As a preliminary to this discussion, the two multiple-fronting structures are spelled out in more detail. For the [+MFS] languages I argue for a hierarchical structure within SpecCP, such that noninitial WHs are adjoined to the right of the specifier position, while the initial one is the head of SpecCP. For the [-MFS] languages, I argue for a structure in which each noninitial WH is adjoined to IP. Among other advantages, these structures make it possible to bring the word order differences between the two types of languages under a version of the split ECP, specifically a requirement that Wh-traces be locally bound at LF.

Sections 2.4 and 2.5 present several factors which, taken together, account for much of the crosslinguistic variation in Wh-constructions. One of these, the Condition of SpecCP Adjunction, holds that languages differ in the levels at which they permit adjunction to SpecCP. Other factors considered are the level(s) at which Wh-movement applies, whether Comp can be intrinsically [+WH], and whether LF Wh-movement is from A or A' positions. Further work may well show some of these to be derivable from more basic parameters.

In closing, I would like to touch briefly on the important issue of learnability. How does a child learning a multiple-Wh-fronting language come to know whether it is a [+MFS] or a [-MFS] language? More generally, how does any child know where its language fits in the range of types available for Wh-constructions, particularly multiple questions?

Following Adams (1984), I assume that the unmarked value for the CSA is that it applies at all levels. The child thus needs positive evidence in the form of multiple-Wh-questions to conclude that a language allows adjunction to SpecCP at LF, and in the form of Wh-island violations or multiple extraction to conclude that it does so at S-structure. Evidence for SpecCP adjunction at PF could include multiple Wh-phrases preceding a complementizer, or perhaps second position clitics following multiple WHs, if cliticization is sensitive to PF. Without such evidence at each level, the language learner assumes adjunction to SpecCP is impossible. This part of the system is thus definitely learnable, and a child presented with evidence for SpecCP adjunction at all three levels would have all the information necessary to know he or she was learning a [+MFS] language.

Similarly, positive evidence is available for the level at which Wh-movement operates. In the absence of overtly moved Wh-phrases the learner assumes no syntactic Wh-movement. In the absence of WH in situ the learner assumes no LF Wh-movement from A positions. And in the absence of a [-MFS] pattern (see below) the learner assumes no LF

movement from A' positions. If both moved and in situ or [-MFS] pattern Wh-phrases are found, the learner concludes that Wh-movement applies at both levels.

By a [-MFS] pattern I mean multiple Wh-fronting with no subject/object word order asymmetry, second position clitics and parentheticals following the first Wh-word, and no multiple extraction. Any evidence from clitics would be especially useful for a learner, since clitics are found in short, high-frequency sentence types. This [-MFS] pattern allows the child to conclude that he or she is learning a language with LF movement from A' positions, i.e. a [-MFS] language.

#### REFERENCES

- Adams, Marianne: 1984, 'Multiple Interrogation in Italian', *The Linguistic Review* 4, 1-27.
- Aoun, Josef: 1985, *A Grammar of Anaphora*, MIT Press, Cambridge, Mass.
- : 1986, *Generalized Binding*, Foris, Dordrecht.
- Aoun, Josef, Norbert Hornstein, David Lightfoot, and Amy Weinberg: 1987, 'Two types of Locality', *Linguistic Inquiry* 18, 537-577.
- Aoun, Josef, Norbert Hornstein, and Dominique Sportiche: 1980, 'Some Aspects of Wide Scope Quantification', *Journal of Linguistic Research* 1, 69-95.
- Aoun, Josef, and Dominique Sportiche: 1983, 'On the Formal Theory of Government', *The Linguistic Review* 3, 211-235.
- Browne, Wayles: 1976, 'Two WH Fronting Rules in Serbo-Croatian', *Južnoslovenski Filolog* 32, 195-203.
- : 1986, *Relative Clauses in Serbo-Croatian in Comparison with English*, Institute of Linguistics, University of Zagreb.
- Chomsky, Noam: 1973, 'Conditions on Transformations' in Stephen R. Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York, pp. 232-286.
- : 1981, *Lectures on Government and Binding*, Foris, Dordrecht.
- : 1986, *Barriers*, MIT Press, Cambridge, Mass.
- Cichocki, Wladislaw: 1983, 'Multiple Wh-questions in Polish: A Two-Comp Analysis', *Toronto Working Papers in Linguistics* 4, 53-71.
- Comorovski, Ileana: 1986, 'Multiple WH Movement in Romanian', *Linguistic Inquiry* 17, 171-177.
- Contreras, Heles: 1986, 'ECP as an S-structure Principle', unpublished ms., University of Washington.
- Davis, Lori J.: 1987, 'Remarks on Government and Proper Government', *Linguistic Inquiry* 18, 311-321.
- Fiengo, Robert: 1980, *Surface Structure*, Harvard University Press, Cambridge, Mass.
- Higginbotham, James and Robert May: 1980, 'Questions, Quantifiers, and Crossing', *The Linguistic Review* 1, 41-80.
- Horvath, Julia: 1981, *Aspects of Hungarian Syntax and the Theory of Grammar*, unpublished Ph.D. dissertation, UCLA.
- Huang, C. T. J.: 1982, 'Move WH in a Language without WH Movement', *The Linguistic Review* 1, 369-416.
- Jaeggli, Osvaldo: 1986, 'Quantification and WH Questions in Spanish', unpublished ms. University of Southern California.
- Kayne, Richard: 1980, 'Extensions of Binding and Case-Marking', *Linguistic Inquiry* 11, 75-96.
- : 1981, 'ECP Extensions', *Linguistic Inquiry* 12, 93-134.



- : 1984, *Connectedness and Binary Branching*, Foris, Dordrecht.
- Koopman, Hilda: 1984, *The Syntax of Verbs*, Foris, Dordrecht.
- Kraskow, Tina: 1987, 'Implications of Multiple Wh-Movement for Wh-island Violations', paper read at LSA annual meeting, San Francisco.
- Lasnik, Howard and Mamoru Saito: 1984, 'On the Nature of Proper Government' *Linguistic Inquiry* 15, 235–289.
- Lie, Svein: 1982, 'Discontinuous Questions and Subjacency in Norwegian', in Elisabet Engdahl and Eva Ejerhed (eds), *Readings on Unbounded Dependencies in Scandinavian Languages*, Almquist and Wiksell, Stockholm, pp. 193–203.
- Maling, Joan M.: 1978, 'An Asymmetry with Respect to WH Islands', *Linguistic Inquiry* 9, 75–89.
- McCloskey, James: 1979, *Transformational Syntax and Model-Theoretic Semantics: A Case Study in Modern Irish*, Reidel, Dordrecht.
- Penčev, Jordan: 1984, *Strož na bulgarskoto izrečenie*, Nauka i Izkustvo, Sofia.
- Pesetsky, David: 1987, 'Wh-in-Situ: Movement and Unselective Binding', in E. Reuland and A. G. B. ter Meulen (eds.), *The Representation of (in)definiteness*, MIT Press, Cambridge, Mass., pp. 98–129.
- Reinhart, Tanja: 1982, 'A Second COMP Position', in A. Belletti, L. Brandi and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar*, Scuola Normale Superiore di Pisa, Pisa, pp. 515–557.
- Rizzi, Luigi: 1982, 'Violations of the Wh Island Constraint in Italian and the Subjacency Condition', in L. Rizzi (ed.), *Issues in Italian Syntax*, Foris, Dordrecht, pp. 49–76.
- Rudin, Catherine: 1982, 'Who what to whom said? An Argument from Bulgarian against Successive Cyclicity', *Proceedings of the Chicago Linguistic Society* 17, 353–360.
- : 1985, 'Multiple WH Movement and the Superiority Condition', paper read at LSA annual meeting, Minneapolis.
- : 1986, *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*, Slavica, Columbus.
- : 1988, 'Multiple Questions in South Slavic, West Slavic, and Romanian', *Slavic and East European Journal* 32, 1–24.
- Toman, Jindřich: 1981, 'Aspects of Multiple WH Movement', in R. May and J. Koster (eds.), *Levels of Syntactic Representation*, Foris, Dordrecht, pp. 293–302.
- : 1986, 'Cliticization from NPs in Czech and Comparable Phenomena in French and Italian', in H. Borer (ed.), *The Syntax of Pronominal Clitics*, Academic Press, New York, pp. 123–145.
- Torrego, Esther: 1984, 'On Inversion in Spanish and some of its Effects', *Linguistic Inquiry* 15, 103–130.
- Travis, Lisa: 1984, *Parameters and Effects of Word Order Variation*, unpublished Ph.D. dissertation, MIT.
- Wachowicz, Kristina A.: 1974, 'Against the Universality of a Single Wh-Question Movement', *Foundations of Language* 11, 155–166.
- Williams, Edwin: 1986, 'A Reassignment of the Functions of LF', *Linguistic Inquiry* 17, 265–299.
- Williams, Joanna Radwanska: 1986a, 'The Structure of Questions, Wh Movement, and Subjacency in Polish', unpublished ms., University of North Carolina.
- : 1986b, 'WH Constructions and Underlying Word Order in Polish', paper read at meeting of American Association of Teachers of Slavic and East European Languages, New York City.

Received 17 September 1986

Revised 22 January 1988

Humanities Division  
Wayne State College  
Wayne, NE 68787, U.S.A.