

MULTIPLE QUESTIONS IN SOUTH SLAVIC, WEST SLAVIC, AND ROMANIAN*

Catherine Rudin, Wayne State College (Nebraska)

Introduction. The fact that some languages allow or require all interrogative words ("K-words")¹ to move to the beginning of the clause in multiple questions has been recognized at least since Wachowicz noted Slavic language examples like those in (1) (1a = her 6, 1b = her 12).

- 1a. **Kto** **čto** **kogda** skazal? (Russian)
 who what when said
 "Who said what when?"
- b. Powiedz mi **kto** **gdzie** mieszka. (Polish)
 tell me who where lives
 "Tell me who lives where."

In addition to the Slavic languages, this pattern is followed by some non-Slavic languages of East Europe, including Romanian and Yiddish. Although multiple questions in several of these languages have been discussed in the generative literature, no agreement has been reached concerning their structure. In traditional grammars the discussion of multiple questions is usually limited to noting the existence of an example or two without further comment. In this article I examine constructions like (1) in Bulgarian, Serbo-Croatian, Polish, Czech, and Romanian, demonstrate that these languages fall into two groups with respect to several aspects of the syntactic behavior of multiple questions, and argue that the differences between the two groups are attributable to a difference in surface constituent structure.

Two general types of analyses have been proposed for multiple questions like those in (1) within the Government-Binding (GB) theory of Chomsky (*Lectures*), the framework I assume here. In the first type, the K-words or phrases are all in Comp (Complementizer position; a distinct syntactic constituent at the beginning of the clause), while in the second type some or all of them are not in Comp, but rather in some S-internal position. These two analyses would give the following structures for sentence (1a), omitting details of the internal structure of Comp and S:

SEEJ, Vol. 32, No. 1 (1988)

↳ The Slavic and East European Journal ¹

The word order and overall appearance of these simple questions is identical in the two languages. In spite of the important differences between Bulgarian and Serbo-Croatian which we will see shortly, many aspects of K-fronting are the same in both. The most important similarity between the two languages, for our purposes, is that multiple K-words are fronted in a number of constructions, including not only main-clause and embedded questions, but also free relatives, concessive clauses, and a construction in which K-words function as indefinite pronouns. One example of each type for each language is given below.

question:

- 5a. **Koj kakvo e kupil?** (B)
 who what has bought
 "Who bought what?"
- b. **Ko šta hoće?** (SC)
 who what wants
 "Who wants what?"

indirect question:

- 6a. **Ne znaem koj kakvo kâde e kupil.** (B)
 not know-1p who what where has bought
 "We don't know who bought what where."
- b. **Ne znam ko šta hoće.** (SC)
 not know-1s who what wants
 "I don't know who wants what."

free relative:

- 7a. **Koj(to) kakvoto e kupil da go vârni vednaga.** (B)
 who what has bought to it return immediately
 "Whoever bought whatever, let him return it immediately."
- b. **Ko šta hoće neka uzme.** (SC)
 who what wants may-he take
 "Whoever wants whatever, let him take it."

concessive free relative:

- 8a. **Vseki se otbivaše, kojto kogato i da mineše.**³(B)
 each refl dropped-in who when and to passed
 "Each dropped in, whoever whenever was passing by."
- b. **Ko god šta hoće, neka uzme.** (SC)
 who ever what wants may-he take
 "Whoever wants whatever, let him take it."

- 11a. *Šte kupim povečeotkolkoto kakvoto ni e dal. (B)
 will buy-1p more than-how-much what to-us has given-3sm
 ("We will buy more than [how much] he gave us what.")
- b. *Kolko kakvi koli! (B)
 how-many what-kind cars
 ("What a lot of what cars!")

In short, in both Bulgarian and Serbo-Croatian, all K elements can always front and all K-constructions can be multiple as long as this does not lead to impossible binding configurations or other problems with semantics or construal. The differences between the two languages which we will see in the next section are not due to any difference in what constructions can be multiple.

1.2. Some Differences between Bulgarian and Serbo-Croatian. Despite the great similarities in the types of K-constructions and particularly multiple-fronting constructions that they allow, there are also important differences between the K-constructions of Bulgarian and Serbo-Croatian. Upon closer inspection the two languages turn out to differ significantly in several ways with respect to the position and behavior of K-words.

One area of difference is the position of clitics relative to the K-words. In Bulgarian the clitic auxiliary *e* must not come between *koj* and *kogo*, while in Serbo-Croatian the corresponding clitic *je* must be between the first and second K-words; when this is violated, the result is an ungrammatical sentence.

- 12a. **Koj kogo e vidjal?** (B)
 who whom has seen
 "Who saw whom?"
- b. ***Koj e kogo vidjal?**
 who has whom seen
- c. **Ko je koga video?** (SC)
 who has whom seen
- d. ***Ko koga je video?**
 who whom has seen

The same is true of other clitics, such as clitic object pronouns: they all follow the K-word group in Bulgarian but split it in Serbo-Croatian.

- 13a. **Koj kakvo ti e dal?** (B)
 who what you has given
 "Who gave you what?"
- b. ***Koj ti e kakvo dal?**

- e. **Kome** je što ko dao?
 f. **Kome** je ko što dao?
 "Who gave what to whom?"

A third difference between the two languages is the relative acceptability of K-words in situ. In Bulgarian a sentence like (18a) is acceptable only as an echo, that is, as an incredulous or uncomprehending response to a statement like **Georgi e vidjal bašta si** "Georgi saw his father." (# marks echo status.) As has been noted by Penčev (66), in order to be interpreted as a "real," nonecho question, a multiple question must have all of its K-words fronted. In Serbo-Croatian, on the other hand, sentences like (18b), with only one of the multiple K-words fronted, or even (18c), with both Ks in situ, appear to be much more normal. The judgments here are quite subtle and variable.⁸ Nonetheless, there is a definite difference in degree of acceptability of such questions between the two languages.

- 18a. #**Koj** e vidjal kogo? (B)
 who has seen whom
 b. (?#)**Ko** je video koga? (SC)
 who has seen whom
 c. (?#)**Video** je ko koga? (SC)
 seen has who whom

A fourth difference is in the possibility for movement of multiple K-words out of an embedded clause ("long movement"). In Bulgarian, all of the K-words in a multiple question must move to a higher clause if the clause in which they originate is noninterrogative (for instance, a *če* 'that' clause); K-words in situ (19b, c; 20b, c) are ungrammatical or interpretable only as echoes, just as they are when in a single clause, and leaving one K in the Comp of the lower clause, as in (20d, e), is also impossible. Blanks in the examples indicate the position from which the K-words presumably moved.

- 19a. **Koj** kâde misliš [če e otišâl ___ ___]? (B)
 who where think-2s that has gone
 "Who do you think (that) went where?"
 b. ***Koj** misliš [če e otišâl ___ kâde]?
 who think-2s that has gone where
 c. ***Kâde** misliš [če e otišâl koj ___]?
 where think-2s that has gone who
 20a. **Ksenija na kogo kakvo** kaza [če šte donese ___ ___]?⁹
 Ksenija to whom what said that will bring-3s
 "What did Ksenija say that (she) would bring to whom?"

Except for this case, however, long movement of multiple K-words is grammatical in questions. Long movement is possible also in both types of multiple free relatives in Bulgarian¹¹ (cf. 7a, 8a); these presumably have the same structure as multiple questions.

- 24a. **Koj kakvoto** misliš će šte iska — —, damu go kupiš. (B)
 who what think-2s that will want to him it buy-2s
 “Whoever you think will want **whatever**, buy it for him.”
 (i.e., for each person, buy what you think that person will want)
- b. Vseki šte se otbie, **kojto kogato** i da se seti će trjabva da mine.
 each will drop in who when and to remember that ought to pass
 “Each will drop in, **whoever whenever** remembers that he ought to pass by.”

In Serbo-Croatian long movement constructions, on the other hand, all but one of the K-words can, and for most speakers must, remain unfronted; a second K-word can move neither to the higher clause nor to the Composition of its own clause. (Examples from Steven Franks, personal communication.)

- 25a. **Ko** želite [da vam **šta** kupi ___]? (SC)
 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?”
- 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

Some speakers do accept sentences like (26a), with multiple K-words moved out of their clause, but even these speakers also accept the same sentence with one K left behind, as in (26b, c).

- 26a. **Ko koga** misliš [da je video ___ ___]? (SC)
 who whom think-2s that has seen
- b. **Ko** misliš [da je **koga** video ___]?
 who think-2s that has **whom** seen
- c. **Koga** misliš [da je **ko** video ___]?
 whom think-2s that has **who** seen

Another argument for constituent status comes from definiteness marking in free relatives. Although each K-word may be separately marked with the definite suffix *-to*, which normally marks relative as opposed to interrogative K-words, one *-to* for the K group is sufficient. This strongly suggests that the two K-words form a unit: *[[koj kakvo]-to]*.

30. **Koj** **kakvoto** iska . . . = **Kojto** **kakvoto** iska . . . (B)
 who what-def wants who-def what-def wants
 “Whoever wants whatever . . .”

In Serbo-Croatian, on the other hand, it is very unlikely that all the K-words are in Comp or even a constituent at all. As we have seen, only one K-word can undergo long (i.e., Comp-to-Comp) movement in Serbo-Croatian, so the argument that all of the K-words must be in Comp in order to move out of the clause does not apply. In addition, if multiple K-words were in Comp in Serbo-Croatian, it would necessitate not only having sentence clitics in Comp, but also some other particles, such as the demonstrative particle *to* in sentence (31).¹³ (Note that this is not the same as the suffix *-to* in the previous paragraph.) There is no reason to believe that these particles and clitics are in Comp, and the analysis on which one and only one K-word is in Comp in Serbo-Croatian accounts for the fact that one and only one of the K-words precedes the clitics.

31. **Tko** je to **koga** udario? (Browne) (SC)
 who has that whom hit
 “Who hit whom?/Who was that hitting whom?”

In short, the MFComp parameter accounts for the differences between Bulgarian and Serbo-Croatian both in the ability of multiple K-words to move long-distance, and in their resistance or openness to intervening material. In addition, the MFComp Parameter may also account for the fact that Bulgarian requires a nominative to precede an accusative, while in Serbo-Croatian the order of K's is free. It is more likely that some general principle will produce strict ordering when all of the K-words form a single constituent, as they do in Bulgarian, than when they do not, as in Serbo-Croatian; the principle involved may be some version of the Empty Category Principle.¹⁴

Only one of the differences I have noted between the two languages is not accounted for by the MFComp Parameter, namely why K in situ is more acceptable in Serbo-Croatian. This is not particularly surprising, since it is well known that fronting of question words can be optional in some languages regardless of the structure of Comp. French, for example, allows exactly one question word to front to Comp, just like English does. However, in French, but not English, this fronting is optional:

32. **Who** did you see? ≠ You saw **who**? (echo)
 Qui as tu vu? = Tu as vu **qui**? (nonecho, “normal” question)

2. The Multiply Filled Comp Parameter in Other Languages. If the dichotomy set up by the MFComp parameter is valid, we should expect it to extend to languages other than the two on which it is based. That is, we should expect to find that other languages with multiple K-fronting will fall into one pattern or the other and not something in between. They should exhibit either the cluster of “Multiply Filled Comp (+MFC)” properties characterizing Bulgarian—long movement of multiple K-words, noninterruption of the K-word sequence, and relatively set word order among the K-words—or else the cluster of “-MFC” properties characteristic of Serbo-Croatian—no long movement of more than one K-word, “second position” clitics following the first K-word, and relatively free K-word order. In addition we should check whether +MFC languages have only echo interpretation of nonfronted K-words, like Bulgarian, while -MFC languages have nonecho interpretation, like Serbo-Croatian—although it is unlikely that this difference between Bulgarian and Serbo-Croatian is actually linked to the Comp Parameter.

These expectations do seem to be borne out. The multiple K-fronting languages that have been discussed in the literature do fall rather neatly into +MFC and -MFC groups with all the relevant properties (except for the difference in acceptability of nonfronted Ks, which we have already noted is probably unrelated). In what follows I look briefly at three more languages which have multiple fronting of K-words, Romanian, Polish, and Czech.

2.1. Romanian. Romanian is very clearly a +MFC language, sharing all of the characteristic features of Bulgarian. I will not discuss the interpretation of the facts in any detail, since it is the same as that given above for Bulgarian, but simply present the relevant data.

First, multiple long distance movement is possible in Romanian (as has been noted previously by Adams and Comorovski); in fact both K-words must be fronted to the matrix clause Comp in questions like those in (35).

35a. Cine cu ce vrea [să mearga — —]?
 who with what wants to go

“Who wants to go how?” (i.e., by what means of transportation)

b. Cine cu cine crezi [că a mers — —]?
 who with whom think-2s that has gone

“Who do you think went with whom?”

Secondly, clitics follow the entire K-word sequence rather than just the first K-word. This is true of clitic pronouns like *le* ‘them’ in (37) as well as the auxiliary *a* ‘has’.

2.2 Polish. Unlike Bulgarian and Romanian, but like Serbo-Croatian, Polish is a -MFC language. That is, in constructions with more than one K-word fronted, only the first one of the group is in Comp while the rest are in some S-internal position. The data presented in this section show that Polish has the complex of properties characteristic of -MFC languages.

First, long movement of multiple K-words is impossible. This is somewhat difficult to test in Polish, since no extraction of even single K-words from a clause is normally possible. However, there are a few "bridge" verbs which do allow long movement of at least certain single K-words. Cichocki (63) gives the following example:

41. **Co** Janek mówił, że studenci czytają?
 what Janek said that students read
 "What did Janek say that the students read?"

But even with these bridge verbs, long movement of multiple K-words is ungrammatical:

42. ***Dlaczego co** Janek mówił, że studenci czytają?
 why what Janek said that students read
 ("What did Janek say the students read why?")

Extraction of a single K-word from subjunctive complements is also possible. Lasnik and Saito give the example *Co Maria chce, żeby Janek kupił?* lit. "What does Maria want that Janek buy." The -MFC hypothesis predicts that similar movement of multiple K-words, for instance, a question like *Co komu Maria chce, żeby Janek kupił?* "What does Maria want Janek to buy for whom?" should be impossible. Some speakers do accept sentences like this, however. For these speakers, presumably, as for the Serbo-Croatian speakers who accept sentences like (26a), the second K-word can, under certain circumstances, be moved out of the clause by cyclic adjunction to successive S nodes; see the discussion of examples 33 and 34 above.

Secondly, clitics split the fronted K-word group. Clitics in Polish can come in various positions, including clause-second, preverbal, and in some speech styles also postverbal, but crucially not inside NP or other nonclausal constituents. Thus, unlike in Serbo-Croatian, a clitic in second position in the sentence always follows the first major constituent, never the first word if that word is part of a larger constituent. The following examples, taken from Toman (296), show the possible positions for clitics in multiple K-constructions.

- 43a. **Kto** się **komu** podoba?
 who refl. to-whom likes
 "Who likes who?"
- b. **Kto** **komu** się podoba?
 who to-whom refl. likes

requires all K-words to be in A' positions at S-structure. The difference between Polish and Serbo-Croatian may be simply that Serbo-Croatian lacks this requirement. The lack of nonecho multiple questions with K in situ does not provide any support for the -MFC analysis of Polish, but it is not inconsistent with it either.

Finally, word order in Polish multiple questions seems to be quite free, although the order NOMINATIVE < ACCUSATIVE < DATIVE is preferred at least by some speakers.¹⁶ Lasnik and Saito (280) give the following example of free order of nominative and accusative K-words:

- 47a. Maria zastanawiała się, kto co przyniesie.
 Maria wondered refl. who what would-bring
 "Maria wondered who would bring what."
- b. Maria zastanawiała się, co kto przyniesie.
 Maria wondered refl. what who would-bring

Polish offers one additional argument for a separate constituent including only the first one of a group of K-words which is not available, to the best of my knowledge, in Serbo-Croatian. It is possible in Polish to have one relative and one interrogative K-word fronted in the same clause, as in (48) (example from Lasnik and Saito, 253). Assuming that a clause is either interrogative or not, and that Comp reflects the \pm interrogative status of its clause,¹⁷ both K-words cannot be in the same Comp in this sentence, since this would mean that the clause is simultaneously interrogative and non-interrogative (relative). However, if *który* is in Comp and *kogo* is not in Comp, but rather in or adjoined to S, as is the case under the -MFC analysis of Polish, there is no problem.

49. Spotkałeś mężczyznę, który kogo zabił?
 met-2s the-man who whom killed
 "Whom did you meet the man who killed?"

A similar conclusion could be drawn also from Polish sentences with a noninterrogative complementizer in Comp in a clause which also contains a fronted interrogative K-word (example from Lasnik and Saito, 281).

49. Maria powiedziała, że kto odwiedza Janka?
 Maria said that who visits Janek
 "Who did Maria say visits Janek?"

Neither type of construction, the one in (48) or the one in (49), is possible in Bulgarian or to the best of my knowledge in Romanian; this is exactly what we would expect given that in these languages all K-words must be in Comp at S-structure.

2.3. Czech. Czech is extremely similar in most respects to Polish, and it too is clearly a -MFC language, with only the first K-word in multiple ques-

or Focus position within S or perhaps adjoined to S, as suggested by Lasnik and Saito for Polish.

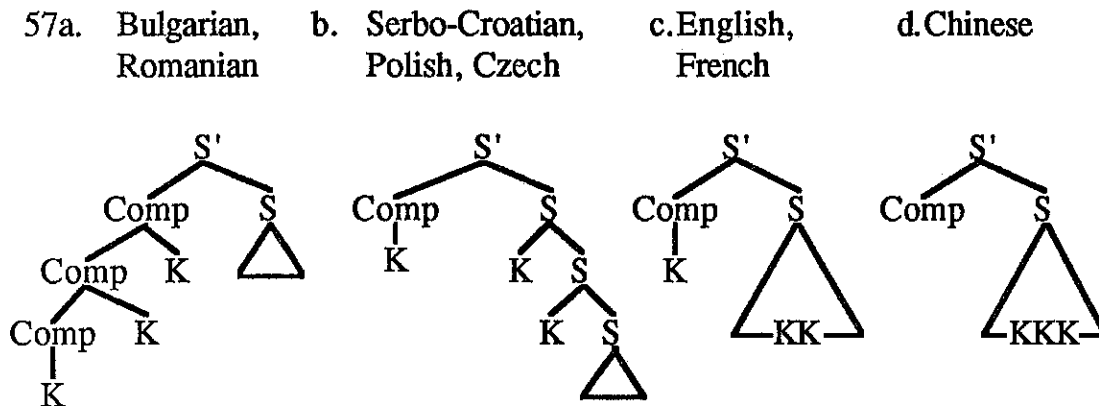
Finally, the order of K-words within the fronted K sequence is relatively loose, as it is in Serbo-Croatian and Polish; the following are all acceptable according to Toman (personal communication):

- 53a. **Kdo kdy koho** pozval, nevím
 who when whom invited I-don't-know
 "Who invited *who when*, I don't know."
- b. **Kdy kdo koho** pozval, nevím
 when who whom invited I-don't-know
- c. **Koho kdy kdo** pozval, nevím.
 whom when who invited I-don't-know

2.4. Conclusions. At this point there are several conclusions that can be drawn. The above survey of several multiple-K-fronting languages bears out the prediction that the Comp Parameter divides languages into two consistent types: Bulgarian and Romanian are opposed to Serbo-Croatian, Polish, and Czech by a complex of properties which can be accounted for by analyzing the former as +Multiply Filled Comp and the latter as -MFC languages. But the survey also makes it clear that certain of these properties are more central than others. In particular, the two most important diagnostic criteria for \pm MFC status are (1) long distance movement of multiple K-words and (2) clitic position within or after the K-word sequence. Not only are these two properties very consistently correlated with each other; they also have an evident causal connection with the MFComp Parameter in the sense that it is clear why the different constituent structures of the K-word sequence result in differences in the possibility of moving it as a unit or splitting it with clitics.

A third property, degree of word order freedom with the K-word group, also seems to correlate quite strongly with the first two, as predicted, although the details of exactly why remain to be worked out. Echo interpretation of K in situ, one of the features which, we have seen, differentiates Bulgarian and Serbo-Croatian, does not turn out to correlate strongly with the other \pm MFC properties. In Bulgarian and Romanian K-movement to Comp (of all K-words) is obligatory in order to form a nonecho question, although it is well known that in some languages, French, for example, such movement is optional. In Polish and Czech fronting of K-words at S-structure is obligatory, but in Serbo-Croatian, another -MFC languages, movement is optional. Whatever factors control differences in the apparent obligatoriness or optionality of K-fronting among languages do not appear to be dependent on the MFComp Parameter.

Assuming, then, the structures (55b) and (54) for the two types of multiple-fronting languages described in this paper, and adding to them the major nonmultiple fronting types, we find that human languages display at least the range of variation shown in (57) for multiple K-constructions at S-structures. In addition to the Slavic and Balkan languages there are languages of the English type, which place only one K-word in Comp at S-structure, and those of the Chinese type, which do not move any of the K-words to Comp except in Logical Form.



The MFComp Parameter separates the +MFC languages (type 57a), which allow adjunction of multiple K-phrases to Comp at S-structure from all three of the other types, which allow it only at LF. Clearly more than this is needed to account for the entire range of possibilities. A more elaborated version of the parameter which recognizes the fact that different languages allow movement of single and multiple K-words to Comp at different levels of the grammar is a major step toward distinguishing the four types of languages:¹⁹

| 58. | single K-word in Comp | multiple K-words in Comp |
|-------------------|--------------------------|-----------------------------|
| A (= type 57a) | S-structure | S-structure |
| B (= type 57b, c) | S-structure | Logical Form |
| C (= type 57d) | Logical Form | Logical Form |

However, a further distinction is needed to separate types (57b) and (57c), which both fall into the same category in chart (58). Both in languages of the English type and in those of the Serbo-Croatian/West Slavic type only one K-word is in Comp at S-structure. The difference between the two groups of languages is that in type (57b) the other interrogative words are moved to nonargument positions at S-structure, while in type (57c) they must remain in situ.

words and, for sake of consistency, even when referring to non-Slavic languages. K-words are boldface in all examples.

- 2 S-structure is a "shallow" level of syntactic structure; for our purposes it is the same as surface structure. For discussion of this and other Government-Binding terms, see Chomsky, *Lectures*.
- 3 Example from Jordan Penčev (personal communication).
- 4 This construction does arguably involve K-movement, although the K-word(s) in it are interpreted as indefinite pronouns, not interrogative or relative forms. See Rudin, *Aspects*, chap. 6, for discussion of the indefinite pronoun construction in Bulgarian.
- 5 Serbo-Croatian does not have an overt K element in unequal comparative clauses, but Bulgarian does:
 - i. Šte kupim poveče dnes, otkolkoto včera.
"We will buy more today than (how much) yesterday."
 Both languages can have an overt K element in equative comparatives:
 - ii. Danas smo kupili koliko i juče. (SC)
"Today we bought as much as (how much) yesterday."
 - iii. Šte kupim tolkova dnes, kolkoto včera. (B)
"We will buy as much today as (how much) yesterday."
- 6 The conditions on ordering of various types of K-phrases are rather complex, but roughly correspond to the template: NOM > ACC > DAT > PP > ADV. For details, see Rudin, *Aspects*, chap. 4. There is some variability across speakers in just how rigid the order of K-words is.
- 7 Examples from Steven Franks (personal communication). One of my consultants finds (14d, f) only marginally possible.
- 8 I have had the same speaker on different occasions give me different judgements of whether these are echos or not.
- 9 In this group of examples *Ksenija* is the sentence initial topic; topic position in Bulgarian is to the left of Comp. Sentence (20e) is grammatical with a reading where *na kogo* is the object of *kaza*, i.e., where it originates in the upper clause. It cannot be taken as the object of *doneše*, as indicated by the blank in the example.
- 10 Absorption, a term due to Higginbotham and May, refers to the fact that multiple questions are usually questions about pairs (or triples). That is, a question like "Who bought what?" is asking for a pairing of people and things: [A bought X], [B bought Y], and so on. In multi-clause questions with K-words in separate Comps this unit pairing does not apply. So "Who knows what we bought?" could be answered "Sarah does."
- 11 But not in the "indefinite" construction shown in (9). This construction, for reasons which I do not understand, is limited to the sequence K*-da-V.
- 12 Although the terminology is different, this analysis of Serbo-Croatian is essentially that proposed in 1976 by Wayles Browne, who notes a number of syntactic differences between the first and second K element in Serbo-Croatian multiple questions.
- 13 On the other hand, as pointed out to me by Steven Franks, the complementizer particle *da* comes after, not within, the K sequence in sentences like (9b). It is not clear to me what the structure is here.
- 14 The in-depth investigation of the Empty Category Principle that would be necessary to support this conjecture is beyond the scope of this paper. See Chomsky (*Barriers*) and other works cited there for discussion of several recent versions of the ECP.
- 15 There is some question about these facts. Joanna R. Williams of the University of North Carolina has claimed in an unpublished paper that it is K in preverbal position that is interpreted as echo, and K in situ as indefinite pronoun, rather than vice versa. But in any case it is clear that in real questions all the K-words are clause initial. Wachowicz stars (46), but indicates that it is bad only in the multiple question reading.
- 16 Williams (personal communication).