

THE COMPLEMENTIZER SYSTEM OF MODERN
STANDARD BULGARIAN

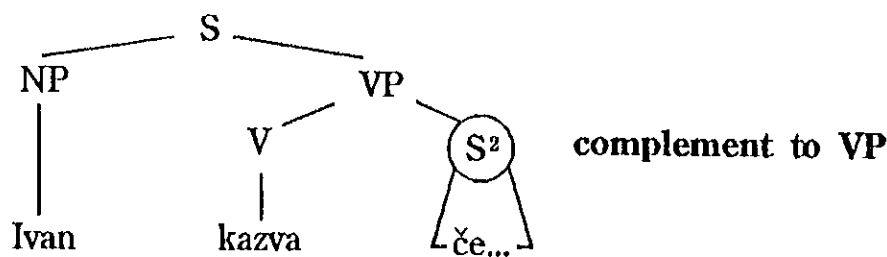
Catherine Rudin

Complementation, that is, the structure of complement clauses and the role of complementizers, has been an area of considerable controversy in theoretical linguistics since the early 1970s. A number of analyses of complementation have been put forward on the basis of English or other languages, and certain features of some of these analyses have been claimed to be universal. The goal of the present paper is to briefly define the term "complementizer" and related notions, then present a (somewhat simplified) analysis of the complementizers of Bulgarian, and describe some of the ways in which Bulgarian data and the analysis presented contribute to our knowledge of complementation in language in general.

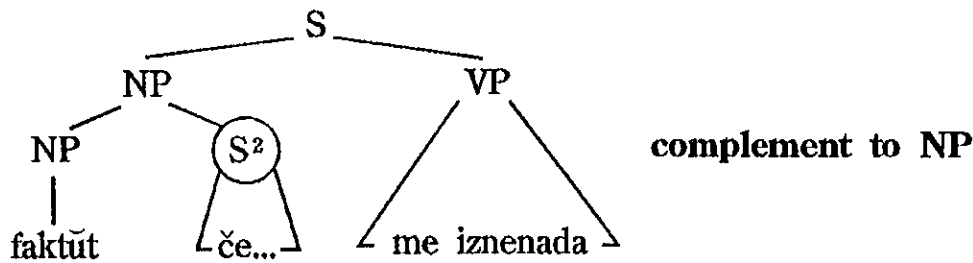
The theoretical framework on which the analysis that I will be presenting is based is a transformational grammar, more or less along the lines of the so-called Revised Extended Standard Theory of Chomsky and his coworkers. The basic features of the analysis are translatable into other theoretical frameworks, however, and I will not be much concerned with questions of theory except in the concluding segment.

Let us begin by defining some terms. First, the definition of COMPLEMENT CLAUSE that I will be using is that *a complement clause is a clause which is subordinated to and selected by some element of the next higher clause*; that is, it is dominated by a Noun Phrase, Adjective Phrase, Verb Phrase, or other constituent of the matrix sentence. Sentences [1a] and [1b] are examples of complement clauses; the underlined clause (the circles "S²" in the tree diagrams) is subordinated to VP in [1a] and to NP in [1b].

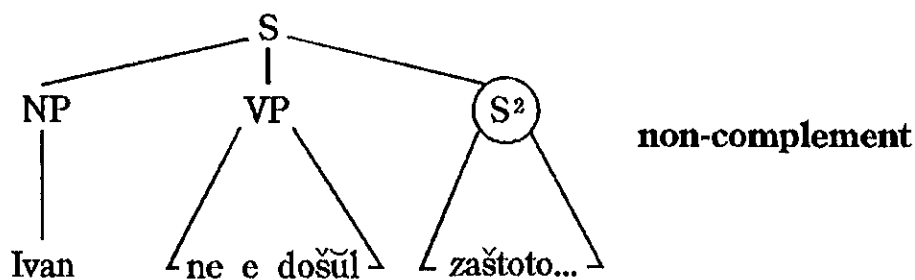
1. a) Ivan kazva če njamaše vreme. (*Ivan pita če njmaše vreme.)



b) Faktūt *če njamaše vreme* me iznenada.



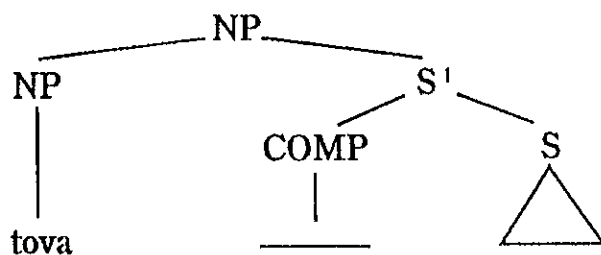
c) Ivan ne e došūl *zaštoto njamaše vreme*.



Non-complement clauses like the one shown in [c] are associated with the entire main clause rather than being dependent on one element of that clause. Complement clauses but not non-complement clauses are selected by the lexical item that dominates them: for instance, *kazvam*, but not *pitam* "takes" a *če* clause in [1a]. It is not the case that *idvam* requires a *zaštoto* clause in [1c], however.

Most subordinate clauses are marked by some kind of conjunction or other linking word, in most languages. In Bulgarian, complement clauses are introduced by the words *če*, *dali*, *deto*, *koj* (and other interrogative proforms), *kojto* (and other relative proforms), and may also contain a particle *da* or *li*. Non-complement clauses can begin with a number of conjunctions of various types, including *zaštoto*, *ako*, *poneže*, and others. A simple test for complement status of clause types is whether a particular linking word can occur in the position labelled COMP in the structure shown in [2]:

2.



Če, dali, deto, koj, and kojto can all occur here, but conjunctions like *ako, zaštoto*, and so on cannot, as the examples, in [3] demonstrate:

3. a) *Tova, če živeeš spokojno, ne e malko.*
 b) *Tova, deto e stanalo v magazina ne e dobro.*
 c) ...spored *tova, dali se otnasja do podloga...*
 d) *Zavisi ot tova, koj pruv e pristignal.*
but f) **Tova, {ako, zaštoto, poneže, etc.}...*

Similarly, clauses beginning with *če, dali, deto, koj, or kojto* can function as sentential subjects or objects (that is, as noun phrases) while those beginning with conjunctions like *ako* cannot: (The only *deto/kojto* clauses that are noun phrases in and of themselves are free relatives, so *kojto* and *deto* are not shown in [4].)

4. a) {*Če, dali, koj*} e došül ne me interesuva. (subj)
 b) Šte vidiš {*če, dali, koj*} e došül. (obj)
but c) *{*Zaštoto, ako, etc.*} e došül njama značenie.
 d) *Šte vidim {*ako, zaštoto, etc.*} e došül.

The interrogative and relative pronouns are similar to *če, dali, and deto* in that they all introduce complement clauses; however, they differ from them in that the interrogative and relative pronouns are *arguments* of the clause. *Koj* represents or "replaces" a subject NP, for instance, while *kogo* represents an object NP, and so on. These words are generally referred to as WH WORDS, while the words which introduce complement clauses and are *not* arguments (that is *če, dali, and deto*, in Bulgarian) are called COMPLEMENTIZERS.

The position in which complementizers occur is called COMP, as we have already seen in tree [2]; in Bulgarian, COMP position follows the TOPIC and precedes the FOCUS position:

5. [TOPIC Ivan] [COMP dali] [FOCUS knjigata] da donese?

This about takes care of the necessary definitions; the notions complement clause, complementizer, COMP position, and WH word have all been introduced, so we are ready to go on to the analysis of complement clause types in Bulgarian. The types mentioned so far are shown in [6].

8. a) *[COMP dali deto] *[COMP dali če] *[COMP če dali]
 b) *[COMP+WH *koj dali*] *[COMP-WH deto kojto]

The analysis which I propose to account for these distributional facts is as follows. WH words are moved into COMP by a rule which places interrogative pronouns in +WH COMP and relative pronouns in -WH COMP. This rule can apply several times if the clause contains more than one WH word, as is the case in example [9].

9. WH MOVEMENT:

Ne e jasno [_S[COMP+WH] *koj iska kakvo.*]

Complementizers are inserted as the surface realization of empty COMP, by a "spelling out" rule like [10]. Since each COMP type is realized as a single specific complementizer, combinations like those in [8a] are avoided, and since only empty COMP (that is, COMP which has not had a WH word placed in it) is spelled out as a complementizer, combinations like those in [8b] also are not produced. This analysis thus rules out all co-occurrences of a complementizer with any other element in COMP, while at the same time allowing more than one WH word in COMP, which is exactly the correct result for Bulgarian.

10. COMP REALIZATION:

COMP - WH → deto

COMP + WH → dali (unless the clause contains *li*)

COMP \emptyset → če

Since WH Movement does not affect non-WH clauses, *če* is always spelled out at least in the embedded complements we are considering here, but the other complementizers are spelled out only if COMP does not contain any WH word or words.

The reader may have noticed that up to now I have been totally ignoring one of the major complement types of Bulgarian, namely, clauses containing *da*:

11. a) *Iskam da otida.* (obj)
 b) *Da otideš ne e vŕzmožno.* (subj)
 c) *Njama po-goljamo ŕstastie ot tova, da stoja blizo do tebe.* (NP-S')

As the sentences in [11] show, *da* clauses are complement clauses; they pass the tests in [3] and [4] with flying colors. At first sight it seems that *da* should simply be added to the list of complementizers, and in fact it has often been treated as subordinating conjunctions similar to *če*. *Da*, however, is *not* a complementizer. First, it is not in COMP position, but instead is attached to the verb, and therefore follows rather than precedes the FOCUS. (See example [5], for instance.) Second, *da* can co-occur with any of the complementizers or WH words. (Again, see [5] for an example with *dali*; others are given in [12] below.) As we have seen, however, none of the complementizers can co-occur with each other or with WH words.

12. a) *Ne znam kakvo da kupja.*
 b) *Njama čovek, kojto (or deto) da razbira tova.*
 c) *Potopete...po takŕv načŕn če nivoto na vodata da se namira mežu belezite.*

Da is thus completely unlike the complementizers in its syntactic behavior, in terms of both position and co-occurrence. Verbs or other matrix elements that seem to subcategorize a *da* clause (for example *iskam*, *zelanie*) I suggest actually select a clause with null COMP and *da* later in the clause, i. e., the structure shown in [13]:

13. [_{S'}[COMP]...da...]

The COMP of this type of clause could perhaps be analyzed as COMP \emptyset ; it is clearly non-WH in any case. Under such an analysis COMP \emptyset would be realized as null when the clause contains *da*, given certain semantic conditions (since constructions like [12c] exist, it could not be null simply whenever the clause contains *da*.) It is not at all clear how to formulate these conditions, however, nor how the COMP Realization rule could have access to semantic conditions if they *were* formulated. It seems more reasonable at least for now to assume that there is a fourth type of COMP, which I label "COMP_{da}" for convenience. The system of COMP types thus includes at least two WH types (+ and -) and two non-WH types (\emptyset and "da").

There may well be a couple of other clause types or subtypes as well. One very interesting problem area is that of comparative clauses. As the sentences in [14a-b] show, true comparatives and equative comparatives both resemble relative clauses in that they contain definite WH words in COMP position; it is tempting to analyze them as a subclass of -WH clauses.

14. a) Ti govoriš *poveče otkolkoto trjabva*.
 b) Ne sŭm tolkova loš, *kolkoto me mislite*.
 c) Ne sŭm tolkova bogat *če (da) ne mislja za pari*.

The -WH complementizer *deto*, however, does not occur in such clauses, as might be expected. True comparatives do not correspond to any complementizer in Bulgarian, while equatives seem to take *če*, if the construction in [14c] can be considered parallel to that in [14b]. Another clause type whose analysis is not clear, that is, it is not clear whether it can be made to fit into the system of COMP types I have outlined or requires setting up a separate COMP type, is exclamative clauses of the type shown in [15].

15. *Kolko e hubava! Kakva hubava roklja!*

Again, these are formally similar to questions, since they contain an indefinite WH word that appears to be in COMP position. It is doubtful, however, that they can be analyzed as +WH clauses.

In conclusion, I would like to very briefly discuss the significance of this analysis of Bulgarian for general linguistic theory. As stated at the beginning of this paper, complementation has been an area of intense research in recent generative theory, and a number of putative universals have been put forward. The Bulgarian material is theoretically interesting insofar as it supports or raises problems for any of these theories, or suggests new approaches to the problem. I will mention just a few such cases.

First, in a very general sense Bulgarian supports the usefulness of the notions "COMP", "Complementizer", and "WH word", since assuming these categories to be available in universal grammar makes possible a very simple and I think satisfying analysis of the basic clause types of the language. (This analysis extends to main clauses as well as complement clauses per se, although I have dealt only with subordinate clauses here.) Several areas of Bulgarian grammar which we have not explored here, for example, restrictions on extraction contexts, are

elegantly describable in terms of the clause types in chart [6] and the notions COMP and WH WORD.

Secondly, Bulgarian seems to support an analysis in which complementizers are inserted or spelled out at the surface, rather than being base generated, and in which WH words are in COMP position. Both of these points are controversial, but both are necessary to the analysis of Bulgarian proposed above, since it is these two factors together which make it possible to account for the co-occurrence restrictions on complementizers, while allowing multiple WH words in COMP.

This brings us to our final point, which is that the analysis of Bulgarian given here directly contradicts the supposed impossibility of multiple WH filling of COMP. It has been suggested (by Noam Chomsky and others) that theoretical considerations should make it impossible for two or more WH words to occupy a single COMP, in any language, but I have claimed that such constructions do occur in Bulgarian. It has not been possible here to defend the analysis of the two WH words in the examples in [7] as both being in COMP; the arguments in favor of this analysis would take us quite far afield. Substantial evidence exists, however, for the correctness of the position assumed here, and if it is correct, it will have quite far-reaching theoretical consequences.

Wayne State University of Nebraska

NOTE

This paper was written for the Third Joint North American-Bulgarian Meeting in 1982 and has not been revised. For a more complete exposition of the analysis presented here, as well as references to the relevant literature, see Catherine Rudin, "Complementizers and WH Constructions in Modern Bulgarian" (Ph.D. Dissertation, Indiana University, 1982). A slightly different analysis of these data, based on updated theoretical assumptions, can be found in Rudin, *Aspects of Bulgarian Syntax* (Columbus: Slavic Publishers, 1986).