

Comparatives and Equatives in Bulgarian and the Balkan Languages*

Catherine Rudin

Certain comparative and equative clauses in several of the Balkan languages share a feature which is of great interest for the study of comparative and equative clauses in general linguistics, namely, they contain a quantifying word or morpheme. This quantifier ("QP" = Quantifier Phrase) appears in boldface italics in the following examples.

1.a. *Rumanian*

Ion e mai înalt decât Maria.
Ion is more tall than-QP Maria

b. *Albanian* (examples from Newmark et al p. 243)

Ai qau më fort sesa unë.
he cried more than-QP I

Ai qau aq sa unë.
he cried as-much QP I

c. *Greek*

Exis perisotera vivlia apo *osa* exo.
you-have more books than QP I-have

Ise *osa* psilos *osa* ego
you-are QP tall QP I

d. *Bulgarian*

Ivan izjade poveče jabülki *otkolkoto* kruši.
Ivan ate more apples than-QP pears

Toj ne e tolkova visok *kolkoto* tebe.
he not is as tall QP you

e. *Macedonian*

Masata e podolga *otkolku* što e široka.
the-table is longer than-QP that is wide

Jas ne sum tolku slab *kolku* nego.
I not am as weak QP him

f. *Serbo-Croatian*

On je (toliko) visok *koliko* i ja.
he is (as) tall QP and I

1984 KK Shangriladze & EW Townsend, ed
Papers for the V Congress of Southeast Eur

The fact that such quantifiers occur is of more than just descriptive interest, because of the support it gives to particular theories of the structure of comparative and equative clauses cross-linguistically and the opportunity it provides for more complete study of these clause types. Bresnan (1973) suggested that comparative clauses in English contain an abstract quantifier phrase, that is, that a sentence like (2a) has an underlying structure similar to (2b):

- 2.a. The table is wider than it is long.
 b. The table is wider than it is [_{QP} x-much] long.

In spite of the fact that no overt quantifier ever appears in such constructions in English or other languages for which abstract QP has been proposed, such as French, this hypothesis has been accepted by many linguists. The abstract QP has a good deal of empirical support, both from the point of view of meaning—everyone would agree that the compared element in a sentence like (2) is semantically or logically quantified—and also for various syntactic reasons. Comparative constructions, for instance, behave like other constructions with a “gap” or missing constituent in that they obey “island” conditions: the comparative in (3b) is just as ungrammatical as the similar question in (3a).

- 3.a. *Who do you know the man who married?
 b. *She has more books than I know a man who has.

One possible explanation for this fact (proposed by Chomsky [1977]) is that both sentences can be analyzed as involving movement of an element out of a “complex NP” (that is, out of the structure [_{NP} . . . [_S x]]); in the case of (3b) the moved element would be the putative abstract quantifier phrase:

- 4.a. *Who* do you know [_{NP} the man [_S who married]]
 b. She has more books than *QP* I know [_{NP} a man [_S who has]]

Borsley (1981) has argued that some equative clauses in Polish are formed by movement of a QP in a way similar to that shown in (4b), while others involve deletion of the QP without movement.

Another type of evidence for the existence of a quantifier phrase in comparative and equative clauses even when it does not show up on the surface comes from restrictions on what other elements may occur in the clause. Pinkham (1982), for example, has pointed out that comparative clauses in French contain the quantified pronoun *en*, and not the non-quantified pro-

nouns *le*, *la*, and so forth, even though the clause contains no overt quantifier:

- 5.a. J'ai plus d'argent que tu *en* as.
 b. *J'ai plus d'argent que tu *l'*as.

In spite of such strong and varied indirect support for the abstract quantifier hypothesis, it is still comforting to find languages like those of the Balkans, in which a quantifier actually appears in the surface form of at least some comparative and equative clauses. The existence of an overt quantifier in sentences like those in (1) above not only provides very concrete support for the quantifier phrase hypothesis, but also makes it possible to investigate various related properties of comparative and equative clauses, such as what position(s) the QP occurs in, and under what circumstances it occurs, with much greater certainty than in languages where the quantifier is not phonetically realized.

A thorough investigation of comparative and equative clauses in all of the Balkan languages would obviously be of great interest; however, in this paper I restrict my attention almost exclusively to Bulgarian. Many of the points made here seem to apply to greater or lesser extent to the other South Slavic and Balkan languages as well, but I do not at present have enough detailed information about them to draw any conclusions.

Comparative constructions in Bulgarian have either the comparative form of an adjective (*po* + adjective) or some other comparative element, such as *poveče* 'more' in the matrix clause, and the comparative clause itself is introduced by either *ot* 'than, from' or *otkolkoto* 'than how much'. Typical examples are given in (6):

- 6.a. Ivan ima *poveče* prijатели *otkolkoto* ima vragove.
 Ivan has more friends than-QP he-has enemies
 b. Toj e *po-umen* *ot* nas.
 he is smarter than us

The *kolkoto* of *otkolkoto* is a quantifier; specifically, it is the definite form of the interrogative word *kolko* 'how much', which is found in both direct and indirect questions, and in exclamations:

- 7.a. *Kolko* struva?
 how-much is-worth 'How much does it cost?'
 b. Ne znam *kolko* vragove ima Ivan.
 not I-know how-much enemies has Ivan
 'I don't know how many enemies Ivan has.'

- c. *Kolko* e umen!
 how-much is smart 'How smart he is!'

This same quantifier is used in equative clauses. Comparisons of equality in which both sides of the equation are expressed generally have *tolkova* 'so much' in the matrix clause and *kolkoto* in the complement. *Tolkova* can, however, be omitted, especially when the comparison has the sense "enough for . . ." Semantically similar result clauses are formed with *tolkova* . . .
če.

- 8.a. Ivan ima *tolkova* prijатели *kolkoto* ima vragove.
 Ivan has so-many friends QP has enemies
- b. Ivan ima prijатели *kolkoto* za dvama.
 Ivan has friends QP for two
 'Ivan has friends enough for two.'
- c. Ivan ima *tolkova* prijатели, *če* gi zabravi.
 Ivan has so-many friends that them he-forgets

These constructions raise at least two interesting points. First, the formal similarity of both equatives and comparatives to relative clauses is extremely relevant to the question of whether comparatives and equatives are "WH" constructions or not, cross-linguistically. (WH constructions are relative clauses, constituent questions, and other clauses which contain a fronted "WH" element: *who*, *which*, *where*, and so on.) Secondly, the conditions under which comparative clauses contain *ot* rather than *otkolkoto* (as in (6b) as opposed to (6a)) are certainly of interest for a description of Bulgarian, and may have some wider interest for linguistic theory as well.

The first of these points is the most easily dealt with. Both comparative and equative clauses in Bulgarian have one obvious feature in common with relative clauses, namely, that they contain a WH word, and furthermore, a definite WH word. *Kolkoto* bears the same relation to the indefinite interrogative *kolko* 'how much?' as relative pronouns like *kojto* 'who' bear to their corresponding interrogatives, in this case *koj* 'who?'. The morphology of *kolkoto* is entirely transparent and entirely parallel to that of relative pronouns, namely WH word + *to*, *-to* being the neuter definite article. There is no doubt that *kolkoto* is a WH word. More significantly, the internal syntax of comparatives and equatives is just like that of relative clauses. In all three types of clauses the WH word must be clause-initial, and may not appear in its apparent underlying position if that is other than clause-initial. (Compare (9b) to (9a), and (9c,d) to (6a), (8a) respectively):

- 9.a. Prijatelite [s *koito* imaš] sa verni.
 the-friends who you-have are true
- b. *Prijatelite [s imaš *koito*] sa verni.
- c. *Ivan ima poveče prijатели ot [s ima *kolkoto* vragove].
 Ivan has more friends than has QP enemies
- d. *Ivan ima tolkova prijатели [s ima *kolkoto* vragove].
 Ivan has so-many friends has QP enemies

Furthermore, the external syntax of all three types of clauses is quite similar as well; all of them occur as complements to noun phrases or adjective phrases. The *tolkova . . . kolkoto* equatives (as in (8a)) in particular fit into the regular pattern of what I have elsewhere called "dummy headed" relatives (Rudin 1982), that is, relative clauses headed by a pro-form of the same type as the following relative pronoun (WH word): *tozi kojto* 'he who', *takŭv kakŭvto* 'the kind which (kind)', *tam kŭdeto* 'there where', *toga-va kogato* 'then when', and so on. This great similarity to relative clauses in both morphological and syntactic structure makes it very attractive to analyze both comparatives and equatives as WH-Movement constructions, at least in Bulgarian, and gives indirect support to the claims of Chomsky and others that such clauses involve WH Movement in languages like English as well. (See above, examples (3) and (4).)

One possible problem with a WH-Movement analysis of Bulgarian comparatives and equatives is that in a sentence like (6a) or (8a), only the determiner of the quantified phrase appears to be moved.¹ The putative derivation of these examples is as in (10):

10. . . . (ot)*kolkoto* ima [NP [det QP] vragove]

This is not the usual pattern for WH Movement, which generally requires the entire major constituent (NP in this case) to be fronted if any part of it is. In English, for example, a relative clause like (11a) is perfectly normal, but its variant (11b) with only the determiner fronted is ungrammatical.

- 11.a. The man [NP [det whose] friends] I met . . .
 b. *The man [det whose] I met [NP — friends] . . .

The whole major constituent normally fronts in Bulgarian as well (. . . [NP [det čiiito] prijатели] posrešt nah . . ., corresponding to (11a), for example); however, unlike in English, there do seem to be some cases when WH Movement can apply to a determiner alone. The embedded question in (12)

not the most usual type of interrogative clause, but it is at least marginally possible, given appropriate stress:

12. Ne znaja *kakva* e kupila [— kniga], no sigurno
 not I-know what-kind she-bought book but surely
 e kupila njakakva.
 she-bought some-kind

It is thus not totally unexpected for a determiner alone to be able to move in constructions like (10) in Bulgarian, although it is surely odd that this is the normal and perhaps even the only type of movement allowed in comparatives and equatives; (13), while perhaps not totally ungrammatical, is much less normal than (10).

- 13.a. ??ima poveče prijateli *otkolkoto* *vragove* ima . . .
 he-has more friends than-QP enemies has
 b. ??ima tolkova prijateli *kolkoto* *vragove* ima . . .
 he-has so-many friends QP enemies has

This objection however is not strong enough to seriously undermine the weight of evidence that both comparatives and equatives in Bulgarian are WH constructions, and are particularly similar to relative clauses.²

The other point I wish to discuss here is the alternation of *ot* and *otkolkoto* in Bulgarian comparatives, that is, the fact that not all comparative clauses contain a quantifier. We have already seen one example of a comparative not containing *kolkoto* (6b), but did not give any indication there of the conditions under which this can occur. The basic generalization is that *ot* rather than *otkolkoto* can³ occur only if the comparative clause is reduced to a single constituent and if that constituent furthermore is either a noun phrase (accusative, as can be seen when it is a pronoun (14b)), an adjective used in the NP-like metacomparative sense (14c), or a deictic adverb (14d).

- 14.a. Georgi e pročel poveče knigi *ot* Ivan.
 Georgi has-read more books than Ivan
 b. Ivan e pročel poveče knigi *ot* tebe.
 Ivan has-read more books than you-acc.
 c. Paltoto e po-skoro sinjo *ot* zeleno.
 the-coat is rather blue than green
 (the coat is more blue than green; closer to blue than green)
 d. Po-dobre e tam *ot* tuk.
 better is there than here

Reasonably enough, these are exactly the constituent types which can occur as objects of a preposition; the structure of all of the above examples is presumably [_{PP} *ot* [<sub>{NP
AdvP}</sub>]]. In all other cases, for instance when the material following *ot/otkolkoto* is or includes a verb (15a), a prepositional phrase (15b), a non-NP adjective (15c), or more than one constituent of any kind (15d), *kolkoto* is required and *ot* alone is impossible.

- 15.a. Tja poveče čete *otkolkoto/*ot* piše.
 she more reads than-QP she-writes
- b. Ivan dava poveče pari na decata *otkolkoto/*ot* na žena si.
 Ivan gives more money to the-children than-QP to wife his
- c. Masata e po-široka *otkolkoto/*ot* dълга.
 the-table is wider than-QP long
- d. Ivan ima poveče kotki *otkolkoto/*ot* Maria (ima) kučeta.
 Ivan has more cats than-QP Maria (has) dogs

The generalization that anything which can be the object of a preposition can occur after *ot* alone, while valid in a rough way, is an oversimplification. In some instances, *otkolkoto* is obligatory even when followed by a single noun phrase. For instance, the comparative in (1d), repeated here as (16a), requires *kolkoto* in spite of the fact that *kruši* 'pears' is perfectly acceptable as the object of a preposition, including *ot*: *sok ot kruši* 'juice of pears/pear juice', for instance. Sentence (16b) is another example which requires *kolkoto*.

- 16.a. Ivan izjade poveče jabŭlki *otkolkoto/*ot* kruši.
 Ivan ate more apples than-QP pears
- b. Vzehme poveče knigi *otkolkoto/*ot* činii.
 we-took more books than-QP dishes

These sentences contrast with superficially similar ones which do allow *ot* alone, like those in (17):

- 17.a. Jabŭlkite sa po-sladki *ot* krušite.
 the-apples are sweeter than the-pears
- b. Ivan pie poveče vino *ot* voda.
 Ivan drinks more wine than water

The difference between these two types seems to have to do not so much with the particular noun phrases involved as with whether the comparison is between NUMBERS OF OBJECTS or not. In cases like those in (16), where

kolkoto is required, the interpretation is that of counting how many of something there are, while (17) and similar cases are concerned with comparing the degree of some property or the quantity of some substance, rather than how numerous a type of object is. In short, *okolto* is required in just those cases where the comparative might answer a question of *how many* in English, rather than *how much* or *how adjective* (how sweet, for example). Verbal aspect and/or definiteness may also be involved in the choice of *ot* vs. *otkolkoto* in comparatives.⁴ In fact, in all of the examples of type (16) that I know of, the noun following *otkolkoto* is a plural indefinite, while the matrix clause verb is perfective. I do not believe that aspect or definiteness correlates directly or causally with obligatoriness of the quantifier; however, they may well have an effect on how likely the "counting objects" reading is.

To summarize the conditions on *ot* and *otkolkoto*, then, I claim that *ot* alone is possible only if (A) it is followed by a single NP or deictic adverb (that is, if *ot* plus the rest of the comparative clause makes a well-formed prepositional phrase, and (B) the sentence is NOT interpreted as comparing the number of objects in two sets.

It is tempting to see condition (A) as related to the distinction between phrasal and clausal comparatives posited by some linguists (see Pinkham [1982] and sources cited there); in other words, to claim that *ot*-comparatives are not comparative CLAUSES at all, but simply prepositional phrases—either generated as such or created out of the "remnant" of a clause from which everything has been deleted except for one constituent. The (surface) structure of the two types of comparatives, the *ot*/phrasal type and the *otkolkoto*/clausal type, would then be something like (18a,b), both possibly but not necessarily being derived from the same underlying form.

- 18.a. Ti jadeš poveče [_{PP} *ot* [_{NP} nego]].
 you eat more than him
- b. Ti jadeš poveče [_S *otkolkoto* toj jade].
 you eat more than-QP he eats

The presence of the quantifier would then correlate with clausal as opposed to phrasal structure of the comparative, and phrasal comparatives would not contain any quantifier (even a silent one) at least in surface structure. This may in fact be true in Bulgarian; however, it does not seem to hold cross-linguistically. Within the Balkans, Rumanian and Albanian have an alternation similar to that in Bulgarian, between comparatives with and without an overt quantifier (Rumanian *decît* vs. *ca*, and Albanian *sesa* vs. *se*; compare the examples in (19) to (1a) and (1b) above).

19.a. Am mai mulți bani *ca* tine.
I-have more money than you

b. Është më shumë për të qarë *se* për të qeshur.
it-is more to cry than to laugh

'It is more fitting to cry than to laugh.' (Newmark et al p. 100)

However, the conditions on when *ca* as opposed to *decît* and *se* as opposed to *sesa* are used appear not to be at all the same as those governing the use of *ot* and *otkolkoto*. This is particularly true for Albanian; note for example that in (19b) *se* occurs in a comparative clause with a verb phrase, a usage which is completely impossible with Bulgarian *ot*. Standard Macedonian also has an alternation between two types of comparatives; with *od* and with *otkolku* (*što*), which to the best of my knowledge follow the same pattern of occurrence as Bulgarian *ot/otkolkoto*. This is of course not surprising considering how closely related the two languages are, and does not constitute good evidence for any sort of universal principle.

The theoretical and cross-linguistic significance of the distribution of *ot* and *otkolkoto* is thus uncertain; however, it is nonetheless an interesting fact for the description of comparative formation within Bulgarian, and one which I believe has not previously been noted. The conditions governing the different types of comparative formation in the other Balkan languages and their relation to the Bulgarian types and to principles of Universal Grammar are unclear at present, but are potentially very interesting and surely deserving of further research.

Miami University

NOTES

* I would like to thank the following people (in no particular order) for helping with the data for this paper: Rada Hanu, Milka Ilić, Manolis Serfiotis, Vasile Munțeanu, and Victor Friedman.

1. This observation is due to Jessie Pinkham.
2. For more detailed description of the similarities between these constructions, see Rudin (1984).
3. *Otkolkoto* can also occur in such sentences in colloquial speech, but *ot* is favored by normative grammarians and is said to be gaining popularity in actual usage as well (Aksela Lazarova, p.c.).
4. A suggestion to this effect was made to me by Howard Aronson.

REFERENCES

- rsley, Robert D.
1981 "WH Movement and unbounded deletion in Polish equatives," *Journal of Linguistics* 17:271-288.
- esnan, Joan
1973 "Syntax of the comparative construction in English," *Linguistic Inquiry* 4.3:275-345.
- omsky, Noam
1977 "On Wh-movement," in Culicover, Wasow, and Akmajian (eds.), *Formal Syntax*. New York: Academic Press.
- wmark, Leonard, Phillip Hubbard, and Peter Prifti
1982 *Standard Albanian*. Stanford: Stanford University Press.
- akham, Jessie E.
1982 *The Formation of Comparative Clauses in French and English*. Ph.D. dissertation, Harvard University, distributed by Indiana University Linguistics Club.
- din, Catherine
1982 *Complementizers and WH Constructions in Modern Bulgarian*. Ph.D. dissertation, Indiana University
- 1984 "Comparing Comparatives," *Papers from the 20th Regional Meeting, Chicago Linguistic Society*. Chicago: CLS.