

Phrasal and Clausal Comparatives: Evidence from Balkan Slavic*

Catherine Rudin
Wayne State College

1 Introduction

It has long been known that comparatives in Balkan languages contain an overt *wh* operator (underlined in (1)),¹ corresponding to the abstract *wh* proposed by Bresnan (1973) for English, as in (2).

(1) Bulgarian

Bill e po-bogat **otkolkoto** e Susan.
Bill is more-rich than.how.much is Susan
'Bill is richer than Susan is.'

(2) Bill is richer than [_{QP} *how much*] Susan is.

When I first discussed this *wh* element (Rudin 1984a,b,c), Bresnan's analysis was fairly new, still quite audacious, and entirely theoretical, so the surface appearance of an overt *wh* word in Bulgarian was an exciting finding. In the last few decades some other languages have also been shown to have overt *wh* quantifiers in comparatives (for instance, some

* Much of this material was presented at the 20th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore, Salt Lake City, UT 4/29/2016. The present version of the paper benefits from comments there as well as at FASL. Thanks are due also to two anonymous reviewers.

¹ Throughout the paper comparative-introducing words are boldfaced; comparative *wh* words/morphemes and complementizers are in addition underlined.

Romance languages; see e.g. Matos & Brito 2008). The concept of a universal comparative operator is now widely accepted, and the simple existence of a *wh* element in (2) is no longer very noteworthy. However, the behavior of such *wh* elements is still well worth investigating. Closer examination of operators like *kolkoto* can elucidate differences among various types of comparatives, and among the various languages with overt comparative operators.

In this paper, I investigate what Bulgarian and Macedonian (and to some extent other Balkan and South Slavic languages with overt comparative *wh*) can tell us about the syntax² of comparative constructions. In particular, I examine comparatives which do and do not have an overt *wh* operator in Balkan Slavic (in Section 3), and ones which do and do not have an overt complementizer (in Section 4), with a view to elucidating some still-unsolved issues in the analysis of phrasal as opposed to clausal comparatives. Section 2 provides some background preliminaries, both on the Balkan data and on different types of comparatives cross linguistically, and Section 5 is the conclusion.

2 Preliminaries

2.1 Some Basic Balkan Data

Examples of comparatives with an overt *wh* operator in Bulgarian and Macedonian as well as several other Balkan and South Slavic languages are given in (3a-8a). In all of the Balkan languages (Bulgarian, Macedonian, Romanian, Albanian, and Greek) the *wh* element is a quantifier ‘how much’; in the neighboring BCS (as in some other non-Balkan Slavic languages (Pancheva 2006)) the *wh* element is a non-quantificational *wh* word meaning ‘what’. All of these languages also have comparatives with no overt *wh* element, but only a preposition, as seen in (3b-8b). In most cases, though not all, the preposition in the (b) version for each language is the same as what precedes the *wh* element in the (a) version. Macedonian *od* and *ot* are allomorphs determined by the voicing of the following segment. Albanian and BCS have more than one choice of preposition; I return to a discussion of BCS *od* vs. *nego* in Section 4.2.

² A rich literature exists on the semantics of comparatives, which I ignore here for lack of space.

(3) Bulgarian

- a. Te sa po-umni **otkolkoto** sme nie.
 they are smarter than.how.much are we
 'They are smarter than we are.'
- b. Te sa po-umni **ot** nas.
 they are smarter than us
 'They are smarter than us.'

(4) Macedonian

- a. Poveќе sakam da rabotam **otkolku** da sedam.
 more like_{1SG} to work_{1SG} than.how.much to sit_{1SG}
 'I prefer to work than to sit.'
- b. Ti si povisok **od** mene.
 you are taller than me
 'You are taller than me.'

(5) Romanian

- a. Am mai mulți bani **decît** ai tu.
 have_{1SG} more much money than.how.much have_{2SG} you
 'I have more money than you have.'
- b. Am mai mulți bani **ca** tine.
 have_{1SG} more much money than you
 'I have more money than you.'

(6) Albanian

- a. S'ka gjë më të bukur **sesa** të shohësh
 neg thing more of beautiful than.how.much to see_{2SG}
 Pogradec.in.
 Pogradec.the
 'There is nothing more beautiful than seeing Pogradec.'
- b. Dashke të dalësh më e zgjuar **nga** unë.
 will to come.out_{2SG} more and smart than me
 'So you want to be smarter than me.'
- c. Kënga qenka më e fortë **se** njeriu.
 song too more and strong than man.the
 'The song is really mightier than man.'

(7) Greek

- a. Ehis perisotera vivlia **apo osa** eho.
 have_{2SG} more books than wh have_{1SG}
 ‘You have more books than I have.’
- b. Ise psiloteros **apo** emena.
 are taller than me
 ‘You are taller than me.’

(8) BCS

- a. Toma ima lepšu sobu **nego što** je ova.
 Toma has nicer room than what is this.one
 ‘Toma has a nicer room than this one is.’
- b. Toma ima lepšu sobu **od** vas.
 Toma has nicer room than you_{GEN}
 ‘Toma has a nicer room than you.’
- c. Toma ima lepšu sobu **nego** vi.
 Toma has nicer room than you_{NOM}
 ‘Toma has a nicer room than you (do).’

At first glance these two types of comparatives seem to correspond to the classical distinction between “clausal” and “phrasal” comparatives, which I introduce in the next subsection: those with the overt *wh* operator look clausal, while those without it look phrasal. This correlation turns out to hold to some extent but not entirely, leading us to a more fine-grained analysis especially of the phrasal type.

2.2 Clausal and Phrasal Comparatives: Theoretical Background

In nearly all studies of comparative constructions, a distinction is made between a clausal type (roughly meaning that what follows *than* includes a verb or other evidence of clausal structure, as in (9)) and a phrasal type (in which what follows *than* is a single constituent, usually a nominal, as in (10)):

(9) “Clausal”

- a. Bill has more shoes than **he needs**/than **Sue does**/than **Sue has**/than **Sue has skirts**/ than **Sue skirts**.
- b. Bill is taller than **the door is high**.

(10) “Phrasal”

- a. Bill has more shoes than **Sue**/than **boots**.
- b. Bill is taller than **6 feet**.

A distinction along these lines is made for instance by Stassen (2006), Pancheva (2006), Merchant (2009), Bhatt & Takahashi (2011), among many others. Some authors use the terminology differently. Bacsikai-Atkari (2014), for instance, reserves the term “phrasal” for comparatives expressed by an inherently Case-marked DP; she thus classifies some Russian and Hungarian comparatives as phrasal but all English ones as clausal.³ But most often the terms are used in a surfacey way: if it looks like a single DP, it is phrasal.

In any case the “clausal” and “phrasal” labels are pre-theoretical and may or may not correlate with actual syntactic analysis. Considerable ink has been spilled over how the superficial form of comparatives relates to more abstract structure, especially for the phrasal type.

Although Bresnan’s proposal was controversial decades ago, the clausal type is now almost universally acknowledged to have a structure something like (11). For the sake of concreteness I show the clause here as TP with a QP operator moved to SpecCP, and *than* as head of PP, but details can differ. What matters is simply that the comparative (the part following *than*) is a full normal clause containing a *wh* operator. Parts may of course be elided under identity to parts of the higher clause, (11b) is an example with nothing missing:

- (11) a. -er [_{PP} *than* [_{CP} QP/op_i [_{TP} ... _i ...]]]
 b. taller [_{PP} *than* [_{CP} QP/op_i [_{TP} the door is _i high]]]

The phrasal type is far less agreed-upon. Superficially in a phrasal comparative the portion following the initial preposition is just a DP, or more rarely some other constituent such as an adverb, typically with the

³ Bacsikai-Atkari appears to use “clausal” to mean “underlyingly clausal”; i.e. she fully buys into the reduced clause analysis in (12) below, writing that “since the clause can be recovered, comparatives formed with *than* are invariably clausal” (3). Correspondingly she considers *than* to be a complementizer in all comparatives.

case appropriate for the object of a preposition in the given language. However, phrasal comparatives are often claimed to have more to them than meets the eye. Several proposed structure types are roughly sketched in (12), again assuming that *than* heads a PP and ignoring other details.⁴

(12) a. **Reduced clause analysis**

[_{PP} than [_{CP} op [_{DP} ~~VP~~]]]

b. **“Direct” PP analysis**

[_{PP} than [_{DP}]]

c. **Small Clause analysis**

[_{PP} than [_{SC} DP Δ]]

The reduced clause type of analysis (Bresnan 1973, McCawley 1988, Pinkham 1985, Bierwisch 1989, among others) posits deletion of all but one constituent of an underlying full clause, in this case a single DP. In other words, this analysis claims that phrasal comparatives have the same underlying structure as clausal ones, but with ellipsis, formalized in the earlier literature as comparative deletion, subdeletion, or stripping and more recently treated under various theories of ellipsis. The “Direct” analysis (e.g. McConnell-Ginet 1973; Brame 1983; Napoli 1983; Hoeksema 1983, 1984) asserts that a phrasal comparative is simply a PP, with a plain NP or DP object, at underlying as well as surface levels of structure.

These two types of analysis have been debated since the 1970s. Each has advantages and drawbacks. The Reduced Clause analysis accounts for the meaning of the construction, specifically for the intuition that parts of the construction are “understood” or recoverable, suggesting deletion or silent elements. *Bill is taller than Sue* clearly

⁴ The category of *than* and similar comparative-introducing words is actually an important question. I assume here and will provide some evidence later that it is P, heading PP, but this is by no means a foregone conclusion. Lechner (2001) argues that the comparative construction headed by *than* is a type of coordinated clause, while many have assumed *than* is a complementizer (for instance, Bacskai-Atkari 2014).

means *Bill is taller than the degree to which Sue is tall*, an intuition easily captured by a formulation like (13).

(13) Bill is taller than [~~x-degree Sue is tall~~]

The Direct (PP) analysis deals more easily with the facts of case marking. *Bill is taller than her* cannot derive directly from the ungrammatical (14); at some level *her* needs to be in a position to receive case from the preposition *than* rather than the normal case for the subject of a clause.

(14) *Bill is taller than [~~x-degree her is tall~~]

The debate between Reduced Clause and Direct types of analyses stalled for a while after the 1980s, as research on comparatives turned almost entirely into semantic channels. In 2006 Pancheva proposed a new idea, the Small Clause analysis (12c), under which a phrasal comparative consists of a DP with an empty predicate which is filled in by copying the matrix predicate. Pancheva (2006, 2010) suggests the Small Clause analysis combines the advantages of both the other analyses, allowing case assignment from outside while still having clausal syntactic characteristics such as a predicate and clausal semantics. It is not entirely clear how small a clause Pancheva has in mind, but it would presumably at least lack the CP layer, making it transparent to case assignment.

It is possible, of course, that more than one of the proposed analyses could be correct in different cases and that not all phrasal comparatives have the same structure. This is what I will conclude, in fact; in what follows I demonstrate that there are both reduced clausal and PP (or small clause) comparatives in Bulgarian and Macedonian, with visibly different morphosyntax.

3 Comparatives With and Without Overt *wh*

3.1 Overt *wh* in Balkan Slavic Clausal Comparatives

With this background, we now return to Balkan Slavic. In Bulgarian and Macedonian, as noted earlier, the clausal/phrasal distinction corresponds partially to the presence or absence of the overt *wh* operator, but not completely.

In Bulgarian, *otkolkoto* (preposition *ot* + *wh* element *kolkoto*) appears in all clausal comparatives; *ot* alone is ungrammatical whenever the comparative is a clause containing a verb, as shown in the (b) versions of (15-18).⁵

- (15) a. Gradinata e po-goljama, **otkolkoto** ni trjabva.
 garden.the is more-big than.how.much us_{DAT} needs
 'The garden is bigger than we need.'
 b. *...**ot** ni trjabva
- (16) a. Bebetata sa mnogo po-umni, **otkolkoto** se
 babies.the are much more-smart than.how.much refl
 smjataše dosega.
 considered till.now
 'Babies are a lot smarter than has been thought till now.'
 b. *... **ot** se smjataše dosega
- (17) a. Da započneš e mnogo po-važno, **otkolkoto** da
 to begin_{2SG} is much more-important than.how.much to
 uspeeš.
 succeed_{2SG}
 'Beginning is much more important than succeeding.'
 b. *... **ot** da uspeeš
- (18) a. Evropejskijat sâjuz se nuždae ot Turcija poveče,
 European.the union refl needs of Turkey more
otkolkoto Ankara ima nužda ot nego.
 than.how.much Ankara has need of it
 'The EU needs Turkey more than Ankara needs it.'
 b. *... **ot** Ankara ima nužda ot nego

The same is true in Macedonian: the *wh* element *kolku* is obligatory with clausal comparatives and the preposition *od* alone is ungrammatical:

⁵ These and most of the other examples in the paper were found online via Google search, and verified by native speakers.

- (19) a. Podobro da umreš na noze **otkolku** da živeeš
 better to die_{2SG} on feet than.how.much to live_{2SG}
 na kolena.
 on knees
 ‘Better to die standing than to live on your knees.’
 b. * ... **od** da živeeš na kolena
- (20) a. Poveće sakam da bidam sama **otkolku** so nego
 more like_{1SG} to be_{1SG} alone than.how.much with him
 da živeam.
 to live_{1SG}
 ‘I’d rather be alone than live with him.’
 b. * ... **od** so nego da živeam
- (21) a. Ovoj poraz boli poveće **otkolku** da
 this defeat hurts more than.how.much to
 zagubevme so 20 poeni razlika.
 lost_{1PL} with 20 points difference
 ‘This defeat hurts more than if we had lost by 20 points.’
 b. * ... **od** da zagubevme ...

So far, it looks like the *wh* operator *kolkoto/kolku* could be a marker of clausal comparatives.

3.2 Overt *wh* in Balkan Slavic Phrasal Comparatives

However, it is not the case that all phrasal comparatives have just the preposition *ot/od*, like those we saw earlier (in (3-8)). In fact, the *wh* element *kolkoto/kolku* does occur in phrasal comparatives. I present the facts of Bulgarian first, in 3.2.1, followed by Macedonian in 3.2.2.

3.2.1 Bulgarian. In Bulgarian, the *wh* operator *kolkoto* is found with several types of phrasal comparatives. It can be followed by a DP (as in (22)) and can also occur with certain other single constituents; namely an Adverb (23) or a PP (24).

(22) **otkolkoto** [DP]

- a. Ženite se nuždajot ot poveče sãn, **ot(kolkoto)**
 women.the refl need_{3PL} of more sleep than-how.much
 m̃žete.
 men.the
 ‘Women need more sleep than men.’
- b. Visokijat xolesterol ubiva poveče **ot(kolkoto)** rakāt.
 high.the cholesterol kills more than-how.much cancer.the
 ‘High cholesterol kills more than cancer.’
- c. Zašto njakoi firmi sa po-bârzi, **ot(kolkoto)** drugi?
 why some firms are more-fast than-how.much others
 ‘Why are some companies faster than others?’
- d. Prodade poveče vurstove **otkolkoto** avtomobili.
 sells more wursts than-how.much automobiles
 ‘It [Volkswagen] sells more sausages than cars.’

(23) **otkolkoto** [Adv]

- a. Po-dobre kâsno, **ot(kolkoto)** nikoga.
 more-good late than-how.much never
 ‘Better late than never.’
- b. Edno i s̃što nešto struva po-skâpo tam,
 one and same thing costs more-expensive there
ot(kolkoto) tuk.
 than-how.much here
 ‘The very same thing costs more there than here.’

(24) **otkolkoto** [PP]

- a. Xorata imat po-goljamo doverie na robotite,
 people.the have more-big trust in robots.the
otkolkoto na samite sebe si.
 than-how.much in only.the self refl
 ‘People have more trust in robots than in themselves.’

- b. William Shakespeare e po-popularen v čužbina,
 William Shakespeare is more-popular in foreign
otkolkoto v rodinata si.
 than.how.much in homeland.the poss
 ‘William Shakespeare is more popular abroad than in his
 homeland.’
- c. ... na po-niski ceni, otkolkoto po vremeto na
 at more-low prices than.how.much at time.the of
 bivšija kmet ...
 former.the mayor
 ‘...at lower prices than at the time of the former mayor...’

In most of these cases, *otkolkoto* can be replaced with just the preposition *ot*, unlike what we saw with clausal comparatives. The parentheses in the DP and Adv examples show *kolkoto* is optional. In general, *ot* can replace *otkolkoto* before DP or Adv but not PP.⁶

3.2.2 Macedonian. The facts of Macedonian are parallel to those of Bulgarian (though it does bring one additional factor into play; see Section 2.3.) The overt comparative operator *kolku* ‘how much’ occurs in all types of phrasal comparatives, just like its Bulgarian cognate *kolkoto*. As shown in (25-27), it occurs with nominal, adverbial, and PP phrasal comparatives. And just like in Bulgarian, the *wh* element is optional with DP and Adverb, though not PP. Comparatives with just the preposition *od* and no *wh* operator are found with DP and adverb, as in Bulgarian. In this set the optionality is shown by *otkolku/od* instead of parentheses, because the preposition has a different allomorph before *kolku*, but the facts are exactly the same of those of Bulgarian.

⁶ It is necessary to qualify this statement with “in general” because of an irrelevant quirk seen in (22d), where *?*ot avtomobili* would be odd at best; *kolkoto* is obligatory with plural indefinite DPs like “cars” where numbers rather than amounts or degrees are being compared; see Rudin (1984b).

odkolku [DP]

- (25) a. Našata kuć a može da primi poveć e gosti **odkolku**
 our.the house can to hold more guests than.how.much
 /od vašata.
 /than yours.the
 ‘Our house can hold more guests than yours.’
- b. Podobro e da imaš ludo dete, **odkolku/od**
 better is to have crazy child than.how.much/than
 lud starec.
 crazy old.man
 ‘It’s better to have a crazy child than a crazy old man.’
- c. signal trieset pati posilen **odkolku/od**
 signal thirty times stronger than.how.much/than
 vselenskiot šum
 universe.the noise
 ‘a signal 30 times stronger than the background space noise’

odkolku [Adv]

- (26) a. Nikogaš ne se čuvstvuvav podobro **odkolku/od**
 never neg refl felt_{1SG} better than.how.much/than
 sega.
 now
 ‘I have never felt better than now.’
- b. Polesno e da započnete utre **odkolku/od** denes
 easier is to begin_{2PL} tomorrow than.how.much today
 ‘It’s easier to start tomorrow than today.’
- c. Mnogu podobro mi e ovde **odkolku/od** tamo.
 much better me is here than.how.much/than there
 ‘I’m much better off here than there.’

odkolku [PP]

- (27) a. Ovaa zagatka e mnogu polesno za decata
 that puzzle is much easier for children.the
odkolku za vozrasnite.
 than.how.much for adults.the
 ‘That puzzle is much easier for children than for adults.’

- b. Podobro e da si vo zatvor **otkolku** na rabota.
 better is to be_{2SG} in jail than.how.much at work
 ‘It’s better to be in jail than at work.’
- c. Imaš poveć e sliki od avtomobilot **otkolku**
 have_{2SG} more pictures of car.the than.how.much
 od tvojata devojka.
 of your.the girl
 ‘You have more pictures of the car than of your girl.’

3.2.3 Brief Excursus on Adverbs. It might seem surprising that adverbs can occur in phrasal comparatives where they appear to be the object of a preposition; however, it is actually quite expected. After all, certain types of adverbs do occur in nounlike usage in many languages, for instance as undoubted objects of prepositions in constructions other than comparatives. In the following Bulgarian examples, the preposition *ot* ‘from,’ is presumably the same as *ot* glossed as ‘than’ in the comparative examples in this paper. It is followed by a DP (‘America’) or an Adverb (‘here’). Notice the English glosses also have a preposition with either a nominal or adverbial object:

- (28) a. Ot [Amerika] li si?
 from America Q are_{2SG}
 ‘Are you from [America]?’
- b. Ot [**tuk**] li si?
 from here Q are_{2SG}
 ‘Are you from [here]?’

Adverbs can also function as subjects of sentences, in both Bulgarian and English:

- (29) Točno **tuk** e ljubimoto mi mjasto.
 exactly here is favorite.the my place
 ‘Right here is my favorite place’

Babby (1974) points out that Russian adverbs not only can occur in comparatives but also take case forms, strongly suggesting they are nominal(ized):⁷

- (30) bol'she obyčnogo
 more usually_{GEN}
 'more than usually'

In fact, *ot/od* in comparatives behaves like any normal preposition, being followed only by a DP or nominalized item including certain types of adverbs. *Kolkoto* is optional exactly where what follows it is something that can normally follow a preposition, giving some support to the idea that *ot/od* (and perhaps other comparative-introducing words cross-linguistically, including *than*) are prepositions.

3.2.4 Case in Phrasal Comparatives. When DP in a phrasal comparative with *otkolkoto/odkolku* is a pronoun (the only situation where Case is visible in Bulgarian and Macedonian, which have lost most of the ancestral Slavic case system), it can be either nominative or objective case. (31) shows this in Bulgarian; Macedonian facts are identical.

- (31) **otkolkoto** nie/nas
 than.how.much we/us
 'than we/us (NOM/ACC)' Bulgarian

The case of the pronoun depends on its role: in (32a) *nie* matches the nominative case of the corresponding subject *te* in the main clause, while in (32b) *nas* has the same accusative case and the same role in its clause as the matrix object *vas*.

- (32) a. Te poveče se bojat sega, **otkolkoto** nie.
 they more refl fear_{3PL} now than.how.much we
 'They are more afraid now than we (are).'
 b. Tova vâlnuva poveče vas, **otkolkoto** nas.
 this excites more you than.how.much us
 'This is more exciting to you than (it is to) us.'

⁷ Thanks to Steven Franks for bringing Babby's work to my attention.

On the other hand with *ot* alone the pronoun is always objective, case being assigned by the preposition.

- (33) **ot** nas ‘than us (ACC)’
 ***ot** nie ‘than we (*NOM)’

This suggests that phrasal comparatives with *otkolkoto* are reduced clauses, with the verb and other material elided, while those with *ot* are not, an idea I formalize in the next subsection.

3.3 Two Types of Phrasal Comparatives

Based on the case facts just discussed, as well as the need for a syntactic position for the *wh* operator, I posit that phrasal comparatives in Balkan Slavic come in two types: CP and non-CP. That is, both the “Reduced Clause” analysis (12a) and one of the other proposed analyses (Direct PP or Small Clause) are instantiated in Bulgarian and Macedonian. This claim is summarized in (34).

- (34) Phrasal comparatives with *wh* are Reduced Clause (CP)
 Phrasal comparatives without *wh* are PP (or SC).

The *wh* element (*kolkoto/kolku*) always indicates an underlying clause (CP); that is, comparatives with the *wh* operator all have full clausal structure, whether their surface form is that of a clausal comparative or a phrasal one in traditional terms. Thus the “clausal” (35a) has the same structure as the “phrasal” (35b/c). The first line in each example is Bulgarian; the second line is the corresponding Macedonian.

- (35) a. **otkolkoto** sme nie [PP ot [CP *wh* [TP sme nie]]]
 otkolku sme nie
 than.how.much are we
 ‘than we are’
 b. **otkolkoto** nie [PP ot [CP *wh* [TP ✕ nie ✕]]]
 otkolku nie
 than.how.much we
 ‘than we’

- c. otkolkoto nas [PP ot [CP wh [TP ~~X~~ nas ~~X~~]]]
otkolku nas
 than.how.much us
 ‘than us’

Conversely, comparatives without the *wh* word are not CPs; instead, their structure is that of a simple PP with a DP, or possibly small clause, object. Pancheva’s (2006) arguments for the small clause analysis are mostly semantic and would take us too far afield. In (36) I simply leave open the option that comparatives without *kolkoto/kolku* could be Pancheva-style small clauses instead of simple PPs; in this case the lack of a CP layer would allow accusative case assignment by the preposition.

- (36) **ot** nas [PP ot/od [DP nas]]
od nas or: [PP ot/od [SC nas Δ]]
 than us
 ‘than us’

In other words, presence of a *wh* operator indicates either a full or reduced CP, while lack of *wh* indicates lack of CP structure. A CP projection is necessary for *wh* movement of the operator, so all *otkolkoto/odkolku* comparatives are CPs, even those which are superficially phrasal. Furthermore, *kolkoto* is obligatory in comparatives which are CPs. The apparent optionality of *kolkoto* before DP and Adv is due to the fact that the DP or Adv can either be a clausal remnant or object of a preposition. The CP projection blocks case assignment by the preposition, so the single DP in a reduced clausal comparative like (35b/c) carries the appropriate case for its role/position within the clause.

4 Comparatives with and without overt complementizer

4.1 Macedonian *što*

Up to this point, Macedonian and Bulgarian behave exactly alike; however, the two languages are not identical in all details. Macedonian adds an additional piece to the comparative puzzle in that it sometimes also allows an overt complementizer to occur, and like the occurrence or lack of the *wh* word *kolkoto/kolku*, this complementizer is diagnostic of a particular type of comparative construction. Bulgarian never has an overt

complementizer in comparatives, but some other languages do, including BCS and Greek; I return to these in the following subsections.

In Macedonian, the complementizer *što* ‘that’ often occurs alongside the operator *kolku*:

- (37) a. Povek e umraat **otkolku** **što** se raġaat.
 more die_{3PL} than.how.much that refl be.born_{3PL}
 ‘More are dying than are being born.’
- b. Deteto e popametno **otkolku** **što** mislat.
 child.the is smarter than.how.much that think_{3PL}
 ‘The child is smarter than they think.’
- c. Počesto čita knigi **otkolku** **što** izleguva so
 more.often reads books than.how.much that goes.out with
 drugarite.
 friends.the
 ‘He reads books more often than he goes out with his friends.’
- d. Stareete pobrzo **otkolku** **što** bi trebalo.
 age_{2PL} faster than.how.much that cond should
 ‘You are getting old faster than you should.’

Comparatives with (*otkolku*) *što* are always full finite clauses. Phrasal comparatives never contain *što*. Furthermore, it seems to make a difference what kind of clause is involved: comparatives consisting of the modal, infinitive-like *da* clause construction, like (19), (20), or (21) never allow *što*, although they do include a finite verb.⁸ Gapped or otherwise reduced clauses do not take the complementizer. A nice example comes from an article about the Chinese military. The headline, (38a), is gapped and has just *otkolku*; but the corresponding sentence in the body of the article, (38b), is a full clause, with *otkolku što*.

⁸ Thanks to Elena Petroska for confirming this judgment. The structure of *da* clauses is a complex issue, which I will not attempt to address here except to note that the **otkolku što ... da* facts appear to favor approaches under which *da* clauses are less than fully finite, e.g. lack a true tense projection.

- (38) a. Tie imaat poveć e vojnici **otkolku** Makedonija
 they have more soldiers than.how.much Macedonia
 žiteli.
 inhabitants
 ‘They have more soldiers than Macedonia (has) inhabitants.’
- b. Tie imaat poveć e od 3 milioni vojnici, što e
 they have more than 3 million soldiers that is
 rečisi milion poveć e **otkolku** **što** Makedonija
 said million more than.how.much that Macedonia
 ima žiteli.
 has inhabitants
 ‘They have more than 3 million soldiers, which is to say a
 million more than Macedonia has inhabitants.’

Some Macedonian grammarians claim a semantic distinction between *otkolku* and *otkolku što*; for instance, the web site “Digitalen Rečnik na Makedonski Jazik” suggests that *otkolku* denotes a preferred choice, while *otkolku što* indicates a comparison. But this clearly does not hold in (38) or numerous other examples. Instead, the distinction is syntactic: *što* indicates a full finite clause as opposed to any type of phrasal comparative, including reduced clausal ones. *Što* is not obligatory in full clausal comparatives, but is overwhelmingly common and apparently preferred, at least by some speakers.

4.2 *Wh* and Complementizer in BCS Comparatives

Macedonian is not alone in allowing a complementizer to appear in some comparative constructions. BCS also admits a complementizer, under conditions intriguingly slightly off kilter from those of Macedonian.⁹

A range of possibilities for forming a comparative in BCS is given in (39-40).

- (39) Clausal
- a. Marijina soba je bolja **nego što** je Ivanova (soba).
 Mary’s room is better than what is Ivan’s room

⁹ I am grateful to Bojan Belić for pointing this out and providing the data in (39-40). There appears to be some dialectal or ideolectal variation in judgments; some speakers find the examples with *li* (39b,c) and (40b) to be questionable or awkward.

- b. Marijina soba je bolja **nego što li** je Ivanova (soba).
- c. Marijina soba je bolja **nego li** je Ivanova (soba).
- d. ?? ... **nego** je Ivanova soba
'Mary's room is better than Ivan's (room) is.'

(40) Phrasal

- a. Marijina soba je bolja **nego** Ivanova (soba).
- b. Marijina soba je bolja **nego li** Ivanova (soba).
- c. * ... **nego što** Ivanova (soba)
- d. * ... **nego što li** Ivanova (soba)
'Mary's room is better than Ivan's (room) is.'

Example (39a) is the typical Balkan pattern for clausal comparatives, with a probably-prepositional introducing word *nego* and *wh* word *što* (by the way, not to be confused with the homophonous Macedonian complementizer *što*). Sentence (39b) shows it is also possible to have the complementizer *li* accompanying the *wh* element, much like the combination of *wh*+complementizer in Macedonian. But unlike in Macedonian,¹⁰ it is also possible for *li* to occur on its own (39c), without the *wh* word. Although clausal comparatives in BCS do not always contain the *wh* element, it seems they need to have either *wh* or complementizer (either *što* or *li* or both); a clausal comparative with just *nego* is marginal (39d).

Conversely, *nego* alone is the norm for phrasal comparatives, which cannot contain *što* (as shown by the starred (40c,d)). However, somewhat unexpectedly *li* can occur (40b). If *li* here is a complementizer, as it undoubtedly is elsewhere in BCS, it suggests that (40b) has a CP structure; that is, it is the reduced clausal type of phrasal comparative.

A further complication is that, in addition to the comparative-introducing word *nego*, BCS comparatives can also be formed with

¹⁰ The string *od što* does occur in Macedonian comparatives, as in (i)

(i) *ke mu donese poveke od što ke dade.*
will him bring more than what will give
'It will bring him more than he gives.'

However, *što* here is the homophonous *wh* word (than [*what* he gives]); this is actually an example of *od*-phrasal comparative. Compare a simple PP example:

(ii) *Od što se plašite?* 'What are you complaining about?

another preposition, *od* (see also (8) above), which does not co-occur with either *što* or *li*, suggesting that comparatives with *od* have no clausal structure but are simply PP. Interestingly, *nego* and *od* comparatives also differ in possible interpretations as well as in the case of their complement.¹¹

- (41) a. *Ženama san treba više od muškaraca.*
 women_{DAT.PL} sleep need more than men_{GEN.PL}
 ‘Women need sleep more than men (do).’
 or ‘Women need sleep more than (they need) men.’
 b. *Ženama san treba više nego muškarcima.*
 women_{DAT.PL} sleep need more than men_{DAT.PL}
 ‘Women need sleep more than men (do).’

The complement of *od* is always Genitive case and the resulting phrasal comparative, *od muškaraca* in (41a), has the familiar ambiguity of phrasal comparatives in many languages, including English: the men here can be interpreted as either subject or object of ‘need’. On the other hand, the complement of *nego* takes a case appropriate to its thematic role within the clause and can only be interpreted as having that role; in (41b) it is Dative, like the corresponding nominal, *ženama*, in the main clause, while in (40) *Ivanova (soba)* is nominative. This reinforces the claim that phrasal comparatives with *od* in BCS are simple PP, while phrasal comparatives with *nego* are (or at least can be) underlyingly clausal.

One loose end is why *što* occurs only in full clausal (39) and not in reduced clausal (40) comparatives. If *li* indicates the presence of a CP projection, we might expect a *wh* operator to be able to occur in phrasal comparatives like (40c,d). Perhaps *li* is not in C in comparatives, but in a lower functional head. *Li* in South Slavic languages marks interrogation and/or focus, and is sometimes analyzed as heading a focus projection instead of CP. In this case the *nego* phrasal comparatives would still be “reduced clausal” but with a somewhat smaller clause, lacking the CP layer, and thus excluding *što*. On the other hand, we know that comparative *wh* operators (and complementizers) cross linguistically are

¹¹ Thanks to Dunja Veselinović for these examples. Compare also (8b/c) which differ in the case of the pronoun *vas* vs. *vi*.

very often silent, and conditions on when they can be overt are idiosyncratic. For the moment I assume *nego* comparatives are CP. BCS thus has the following types of comparatives:

- (42) a. Full Clausal: [_{PP} *nego* [_{CP} (*što*) [_{C'} (*li*) [DP]]]]
 b. Reduced Clausal: [_{PP} *nego* [_{CP} Ø [_{C'} (*li*) [DP]]]]
 c. Direct PP: [_{PP} *od* [DP]]

Although the facts are not quite the same as in Bulgarian and Macedonian, once again we find clear evidence for two types of phrasal comparatives, one which has clausal structure and one which is just PP.

4.3 *Parallels in Other Languages*

Without going into detail, it is worth mentioning that other languages possess similar facts to those presented for Balkan Slavic; for a fuller understanding of clausal and phrasal comparatives much more cross linguistic data should be taken into account. I briefly mention just a few cases here. Among the Balkan languages, Greek also allows a complementizer in comparatives; similar to Macedonian and BCS, but with a twist. In Greek the complementizer always occurs on its own, without an accompanying *wh* word, and apparently occurs only in clausal comparatives, like (43b-c). (Examples from Merchant 2009:135f)

- (43) a. I Maria pezi kithara kalitera **apo** ton Gianni.
 the Maria plays guitar better than the Giannis
 'Maria plays guitar better than Giannis.' (phrasal)
 b. I Maria pezi kithara kalitera **ap'oti** pezi kithara o
 the Maria plays guitar better than.that plays guitar the
 Giannis.
 Giannis
 'Maria plays guitar better than Giannis plays guitar.' (clausal)
 c. Eparhoun perisoteres evdomades se ena hrono **apo** oti
 there.are more weeks in a year than that
 eparhoun meres s'ena mina.
 there.are weeks in.a month
 'There are more weeks in a year than there are days in a month.'

Albanian, like BCS, has two different comparative prepositions with differing properties. Many Slavic languages permit overt *wh* expressions in comparatives, as pointed out by Pancheva (2006), but they differ in which *wh* words appear and under what conditions. In addition to those discussed by Pancheva, the following were mentioned by FASL audience members:¹² Croatian equal comparatives can have *koliko* ‘how much’ instead of the *što* seen in unequal comparatives. And Slovene uses *kod* ‘how’ in phrasal comparatives:

- (44) Marko je veći kod Bojan.
 Marko is taller how Bojan
 ‘Marko is taller than Bojan.’

Finally, consider colloquial English, which also allows overt *wh* in some comparatives.

- (45) a. She’s taller than **what** I am.
 b. Their science requirement is a lot less than **what** ours was
 when I went to college.

This use of *what* is limited to clausal comparatives, and is impossible in all phrasal ones:

- (46) a. *She’s taller than **what** me/I.
 b. *Their science requirement is a lot less than **what** the old
 requirement.

This looks like BCS *što* and different from Bulgarian/Macedonian *kolkoto/kolku*, which do occur in the reduced-clause type of phrasal comparatives, as we’ve seen. Clearly different languages have different conditions on when a *wh* element and/or a complementizer can be overt. The reason for these differences remains opaque, but the fact is well established that comparatives, including some phrasal comparatives, can in principle have the full clausal structure including complete CP layer with *wh* landing site and complementizer head, while others do not.

¹² Thanks to Martina Gracanin-Yuksek for the Croatian comment and Adrian Stegovac for the Slovene.

5 Conclusion

The main conclusion of this paper is that phrasal comparatives are not all the same. The Balkan Slavic languages, Bulgarian and Macedonian, provide clear evidence that there are at least two distinct types of what have been traditionally called “phrasal” comparatives, one of which actually has the structure of a clause.

The presence of an overt *wh* operator, *kolkoto* or *kolku*, indicates that the comparative is a CP even if its superficial form is that of a phrasal comparative; it has clausal architecture including a left-peripheral position containing the *wh* operator. Lack of a *wh* operator characterizes comparatives which are simply PP at all levels of structure (or possibly small clause). In Macedonian, the complementizer *što* further marks full clauses as opposed to those with elided elements, including the modal *da* clauses which perhaps are less than fully clausal in some sense. Overt morphology clearly identifies several different types of comparatives, summarized in table (47), which are less easily separated in languages outside the Balkans.

(47)

Traditional label	Type	Bulgarian	Macedonian	syntactic structure
“clausal”	full clause	otkolkoto	odkolku (što)	clause (CP)
“phrasal”	reduced clause	otkolkoto	odkolku	
	PP	ot	od	underlying PP

Along the way we have also seen some evidence that comparatives are PP, at least in Balkan Slavic, where the preposition *ot/od* takes the same kinds of objects as other prepositions: DP, (nominalized) Adverb, or (nominalized) clause. And finally we have seen that languages in and

out of the Balkans have a range of situations in which overt *wh* operators and/or complementizers can appear, which provide an opportunity for much work to come.

References

- Babby, Leonard. 1974. Towards a Formal Theory of “Part of Speech”. In Richard D. Brecht and Catherine W. Chvany (eds.), *Slavic Transformational Syntax* (Michigan Slavic Materials 10). Ann Arbor: The University of Michigan. 151–181.
- Bacsikai-Atkari, Julia. 2014. *The Syntax of Comparative Constructions: Operators, Ellipsis Phenomena and Functional Left Peripheries*. Doctoral Dissertation, University of Potsdam.
- Bhatt, Rajesh & Shoichi Takahashi. 2011. Reduced and Unreduced Phrasal Comparatives. *Natural Language and Linguistic Theory* 29(3): 581–620.
- Bierwisch, Manfred. 1989. The Semantics of Gradation. In M. Bierwisch and E. Lang (eds.), *Dimensional Adjectives*. Berlin/Heidelberg: Springer Verlag. 71–237.
- Brame, Michael. 1983. Ungrammatical Notes 4: Smarter Than Me. *Linguistic Inquiry* 12: 323–328.
- Bresnan, Joan. 1973. Syntax of the Comparative Clause Construction in English. *Linguistic Inquiry* 4: 275–343.
- Hoeksema, Jacob. 1983. Negative Polarity and the Comparative. *Natural Language and Linguistic Theory* 1: 403–434.
- Hoeksema, Jan. 1984. To be Continued: The Story of the Comparative. *Journal of Semantics* 3, 93–107.
- Lechner, Winfried. 2001. Reduced and Phrasal Comparatives. *Natural Language and Linguistic Theory* 19.4: 683–735.
- McCawley, James. 1988. *The Syntactic Phenomena of English*. Chicago: University of Chicago Press.
- Matos, Gabriela and Ana Brito. 2008. Comparative Clauses and Cross Linguistic Variation: A Syntactic Approach. In O. Bonami and P. Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics* 7: 307–329.
- McConnell-Ginet, Sally. 1973. *Comparative Constructions in English: A Syntactic and Semantic Analysis*. Ph.D. Dissertation, University of

Rochester

- Merchant, Jason. 2009. Phrasal and Clausal Comparatives in Greek and the Abstractness of Syntax. *Journal of Greek Linguistics* 9(1): 134–164.
- Napoli, Donna Jo. 1983. Comparative Ellipsis: A Phrase Structure Account. *Linguistic Inquiry* 14: 675–694.
- Pancheva, Roumyana. 2006. Phrasal and Clausal Comparatives in Slavic. In *Proceedings of FASL 14*, 134–165. Ann Arbor, MI: Michigan Slavic Publications.
- Pancheva, Roumyana. 2010. More Students Attended FASL than CONSOLE. In *Proceedings of FASL 18*, 383–400. Ann Arbor, MI: Michigan Slavic Publications.
- Pinkham, Jessie. 1985. *The Formation of Comparative Clauses in French and English*. New York: Garland.
- Rudin, Catherine. 1984a. Comparing Comparatives. In Drogo, Mishra, and Testen (eds.), *CLS 20, Papers from the 20th Regional Meeting*, 334–343. Chicago: Chicago Linguistic Society.
- Rudin, Catherine. 1984b. Comparatives and Equatives in Bulgarian and the Balkan Languages. In K.K. Shangriladze and E.W. Townsend (eds.), *Papers for the V Congress of Southeast European Studies*, 328–337. Columbus: Slavica.
- Rudin, Catherine. 1984c. More Balkan Comparatives. Presented at American Association of Teachers of Slavic and East European Languages annual meeting, Washington D.C.
- Stassen, Leon. 2006. Comparative Constructions. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 686–690. Amsterdam: Elsevier.

carudin1@wsc.edu