

RESTRICTIVE RELATIVE CLAUSES IN STANDARD MACEDONIAN AND STANDARD BULGARIAN

Eleni Bužarovska

University of Skopje/ Research Centre for Areal Linguistics, MANU, Skopje

elenibuzarovska@mt.net.mk

Abstract

Both Bulgarian and Macedonian belong to Balkan Slavic languages that were subject to the convergent process of balkanization. As a result, certain syntactic patterns from Slavic heritage underwent structural changes adjusting to common Balkan models. One such change occurred in the syntactic patterning of restrictive relative clause formation. The synthetic Slavic model of relative pronoun clause linkage was enriched with a new one: analytic accommodation of the relative clause to the relativized NP (head N) in the main clause by means of a morphologically discontinuous relativizer: an indeclinable pronoun and an anaphoric clitic. Following Comrie's classification (1989) of relativisation strategies responsible for typological variation among different languages, the author compares the syntactic models of relative clause linkage in Macedonian and Bulgarian. The statistical analysis of relative clauses in this paper shows that these languages make use of the same set of strategies, but differ with respect to the degree of their application.

1. Introduction

The aim of this paper is to describe restrictive relative clauses (RCs) in Macedonian (Mac) and compare them both statistically and structurally to their Bulgarian (Bul) counterparts. The description of RCs is based on a semantic-functionalist approach. From a semantic perspective, RCs use presupposed information to restrict the range of potential referents and thus identify the referent of the relativized or head N in the matrix clause (cf. Comrie 1989:139, Givón 1990:647). Thus, formally, a RC consists of a head nominal (head N) and a restrictive relative clause. On a functional plane, the restrictive RC, being an additional member of the head N, modifies the head N by conveying some information about the referent of this NP. This broadly formulated functional definition of RCs makes it possible to treat as relative clauses purpose modifying *da*-clauses (1) and headless RCs as in a proverb (2)¹:

(1) *Baram žena da čuva deca.* Mac
search.1Sg.Pres woman Subj.Mark keep.3Sg.Pres children
'I am looking for a woman to babysit children.'

(2) *Koj pee, zlo ne misli.* Mac
who.M.Sg sing.3Sg.Pres evil Neg think.3Sg.Pres

¹ In contrast to prototypical finite RCs, *da*-relatives do not convey presupposed information (cf. Browne 1986). More on subjunctive relatives in Balkan languages see Bužarovska (2004) and Bužarovska and Tomić (this volume). On the derivation of headless relative clauses see Topolińska (2001).

‘The one who sings, means well.’

Other clausal modifiers can also be covered with the above functional definition. If relativisation is seen as a syntactic strategy that serves to enrich the predicative content of the main clause by providing additional information of its term (argument), then a non-term, such as an adverbial modifier, can be eligible for relativisation as well. In that case a RC functions not only as a modifier of a participant (coded by an NP), but also of a modifier of the event itself. It modifies a parameter of the event which is usually encoded by an adverbial. Depending on what kind of adverbials these RCs modify, they may be formalized as locative, manner, instrumental, etc. adverbial clauses. Similar to the other prototypical relative clauses, adverbial RCs are linked to the main clause immediately after the modified head constituent. These relatives assume the status of subordinate clauses on a sentential level, and contrast with the nominal relative clauses that are incorporated into the head N. In view of the fact that a nominal subordinate RC becomes part of the main clause via NP incorporation, (nominal) relativisation can be classified as a hypotactic operation.

The discussion in this paper is focused on the description of nominal RCs and the relativisation strategies in both languages. This is achieved by filtering out the significant typological properties of RCs that are shared by both languages. The most important typological parameter that determines the choice of the relativisation strategy is the encoding of the role of the head N in the embedded clause (cf. Comrie 1989:147). The following three strategies may be used in both languages: (a) relative pronoun strategy involving the use of declinable pronouns as in (3), (b) mixed analytical strategy that consists of (short) pronoun retention and gapping; it involves the use of the indeclinable pronoun in association with a clitic (4), and (c) gapping that employs only the indeclinable pronoun, often accompanied by a “long” pronoun retention (5).²

(3) *Čovekot koj vlez e moj sosed. Mac*
man+the.M.Sg who.M.Sg come.3Sg.Aor be.3Sg.Pres my.M.Sg neighbour
‘The man who came in is my neighbour.’

² According to Comrie (1989:148-152) the relative pronoun type is characteristic of European languages, such as Russian (i):

(i) *deвушка, которaja prišla* ‘the girl that came’

The retention pronoun type is illustrated by the Hausa example where the pronoun in the RC (*ya*) is an anaphor of the head N:

(ii) *dōkìn dà ya mutùl* ‘the horse that died’

The gap type does not provide any overt indication of the role of the head within the matrix. In other words there is no clitic or other device to carry this information, as in the English *that*-clause (iii):

(iii) *the man that gave the book to the girl*

However, there is a mixed pronoun-retention and gap type as in the Czech example:

(iv) *muž, co to go děvče uhodilol* ‘the man that that girl hit’.

(4) *Čovekot što go sretnavme e moj sosed.* Mac
 man+the.M.Sg that 3Sg.Acc.Cl meet.1Pl.Aor be.3Sg.Pres my.M.Sg neighbour
 ‘The man whom we met is my neighbour.’

(5) *Stignavme do rekata što pievme voda od nea.* (coll) Mac
 reach.1Pl.Aor to river+the.F.Sg that drink.1Pl.Imperf water from she.Acc
 ‘We reached the river that we had drunk water from.’

The corpus-based analysis of RCs in Mac and Bul employed in this paper is aimed at reaching some cross-linguistic generalizations on the basis of the inventory and the distribution of the types of relativisation as defined in Comrie (1989). The description of Mac RCs is based on corpus compiled from journalistic and contemporary prose.³ The examples marked by (*coll*) come from the internet, while the unmarked examples belong to everyday use. The literary prose examples reflect the situation that is characteristic of the standard language: these examples were taken from works whose authors represent three consecutive generations of writers in the last forty years. The obtained results from the analysis of Mac RCs are compared to the the results from the analysis of the Bul RCs based on a corpus of restrictive relative clauses compiled from Bul contemporary literary sources and from newspapers.⁴

2. Semantic-functional classification of RCs

All RCs must fulfill the semantic condition of relativisation: to have a joint constituent in both propositions. Syntactically, the joint constituent is formalized as an element modified by the RC. The modified element or the head may be an obligatory constituent of a predication (a term) or an optional one (non-term). Hence, the semantic classification of RCs is based on whether this joint constiutuent is argument or non-argument. Accordingly, two types of RCs are distinguished: argument and non-argument modifying RCs.

2.1 Argument modification

When a proposition modifies an argument in another proposition it is formalized as an NP with an embedded RC. The modified argument expressed by a head noun (head N) is co-referential with the argument in the second proposition. The co-reference relation imposes a syntactic operation of deletion of the second argument. The deleted argument is represented by a gap (\emptyset) in the second clause. The integration of the second clause into

³ The Mac corpus consists of 200 examples from the newspapers and journals and 200 examples from Mac contemporary prose. The newspapers/journals included: Vest 29.1.05 (V1), 2.02.05 (V2), 3.02.05 (V3); Dnevnik 30/1/05 (D1), 31/1/05 (D2), 1/2/05 (D3); Forum 149 (F1), 152 (F2), and 156 (F3); and Kapital 27/1/05 (K). The literary workd included: Slavko Janevski’s (J) “Omarnini”, Dimitar Solev’s (DS) “Zora zad agolot” and Goce Smilevski’s (S) “Razgovor so Spinoza”.

⁴ The Bul literary corpus was obtained from two books: “Esenni raskazi” by Ivajlo Petrov (IP), (Bălgarski pisatel, Sofija 1978), and “Ljubov i smărt po krivite kruši” by Kristin Dimitrova (KD) (Obsidijan, Sofija 2004). Two more books were included in the corpus with over 150 examples for illustration of a particular relative clause type, but were not used for statistical purposes: “Gospodin Nikoj” by Bogomil Rajnov (BR) (Bălgarski pisatel, Sofija 1971) and “Cenata” by Marko Semov (MS), (Bălgarski bestseler, Sofija 2004).

the first is performed by means of a relative pronoun or a relative operator (RO), which is placed in a clause-initial position. The second clause, headed by RO and containing the deleted coreferential constituent (\emptyset) is embedded in the position adjacent to the head N. The gap controls the RO, so that the RO agrees with the deleted NP in categorial features of gender, number and case.

2.2 *Non-argument modification*

When the second proposition modifies an optional constituent of a predication it is formalized as an adverbial NP or a PP that heads an embedded RC. It is quite common for the adverbial antecedent not to be realized in the surface structure when it is inferrable from the context. Adverbial ROs co-refer with the adverbial modifiers of location, time, manner, instrument, reason, comparison and other non-terms that convey circumstantial information about the main event. Using the criterion of argument vs. non-argument modification, RCs may be classified into two major groups: nominal RCs: subject, object and possessive RCs and adverbial RCs: locative, temporal, manner, instrumental, reason, comparison RCs. The first three adverbial RCs make use of two syntactic types of ROs: adverbial ROs *per se* and prepositional ROs, the latter being composed of a preposition and a declinable relative pronoun. The rest of the adverbial RCs employ prepositional ROs as clause linkage devices.

The discussion in this paper is devoted to nominal RCs. They are characterized by the following syntactic properties: (a) co-reference between the head N and an NP in the dependent clause, with both NPs encoding the joint argument; (b) substitution of the dependent co-referential NP with a relative pronoun; and (c) fronting of the relative pronoun to the clause initial position immediately adjacent to the head N, which results in an empty position (gap) of the moved relative pronoun.

3. Means and strategies of relativisation

RCs in Mac and Bul make use of two types of ROs: relative pronouns and relative adverbs with locative, temporal and manner meaning. The focus of our interest are relative pronouns in both languages. It is worth pointing out that the declinable pronouns *koj/kojto* and the indeclinable *što/deto* are historically related to interrogative pronouns.⁵

The declinable adjectival ROs are recruited for production of a synthetic paradigm used for the declinable relativisation strategy. In Mac and Bul the declinable strategy employs variable relative pronouns (*koj, kojšto*) and (*kojto*), respectively. They carry the same categorial features of case, gender and number as the co-referential deleted NP. The second type of relativisation strategy is based on an analytic paradigm composed of *relativum generale*, the petrified invariable pronoun *što/deto* and a demonstrative anaphoric clitic (cf. Topolińska 2003:307).⁶ The clitic (CI) marks the relation of

⁵ The *relativum generale* *što* has developed from the interrogative *što* in Old Church Slavonic (Topolińska 1996:86). Browne (1986:69) argues that *što* is particularly well-suited for the introduction of RCs, due to its factive nature. It belongs to the COMP category because it has relative and subordinating function, so it is both a complementizer and relativizer (ibid 84).

⁶ It is interesting to note that in Mac translation of *Trlis Evangelarium* written in 1861 only *što* and *deto* were used as relativizers (Topolińska 1997:381). In Bul, *što* is considered to be the second indeclinable pronoun used in literary language, though not found in the analyzed corpus. Pašov (1989:98) claims that *što*

agreement between the head N and the deleted NP by coding the categorial information of the latter. The two languages differ with respect to the presence of the clitic in oblique relativisation. With a definite NP, the clitic is obligatory in Mac and optional in Bul with both declinable and indeclinable ROs. This means that Mac object RCs employ only RO+Cl relativisation strategy, in contrast to Bul, where both RO-Cl and RO+Cl strategies are used. Penčev (1984:116) notes that clitics may occur with declinable ROs in long and complex relative clauses.

4. Semantic structure of ROs

It is important to point out that a relative pronoun performs a double function in a RC: it substitutes the deleted constituent and links the RC to the main clause. Hence, the semantic structure of a RO consists of two elements: (a) an anaphoric constituent that co-refers with the joint argument and (b) a subordinating operator that links the RC with the main clause.

The subordinator is different from the relativizer in that the latter has developed from the former as a means of syntactic accommodation. It is in fact a syntactic operator that transforms the interrogative pronouns into relative (Topolińska 1996:86). In Macedonian, the RO *što* plays the syntactic role of a subordinator, while in Bulgarian the particle *-to* performs the same relativizing, subordinating function. In addition, declinable ROs are characterized by structural iconicity because their morphology reflect the semantic structure, given in (6) and (7):

(6) *Kojto* > *koj* (anaphoric) + *to* (subordinator) Bul

(7) *Kojšto* > *koj* (anaphoric) + *što* (subordinator) Mac

It seems that the presence of the subordinating operator signals an increased force of the subordination mechanism. We can hypothesize that the duality of ROs accounts for the occurrence of *koj* or *kojšto* in appositive RCs: they prefer to be integrated with their heads through bi-componential declinable ROs (*kojšto*) or just with their anaphoric parts (*koj*). In appositive RCs, the lack of true subordination between the head and the RC demands the presence of the anaphoric operator, which blocks the reduction of the declinable RO to its subordinating constituent. On the other hand, restrictive RCs tend to use mono-componential indeclinable ROs for their embedding. In restrictive RCs the RO is often reduced to its subordinator which increases its properties as a complement.⁷

is characteristic of formal written register as compared to the more common *kojto*. The reverse use characterises Mac RCs: *što* is more colloquial than *kojšto*.

⁷ Bul *relativum generale deto* has not derived from *kojto* but from *kădeto*, a calque modelled after the Greek *pu*. The original locative sense of *deto* is present in (i), but has changed into a non-locative, temporal sense in (ii), (see Rudin 1986:133):

(i) *Plažata deto često pluvavme*/'The beach where we often swam.'

(ii) *Pomniš li deto se vidjahme za prvi път?*/'Do you remember when we first met?'

5. Structural similarities of nominal RCs

Mac and in Bul relative clauses share a number of common properties, due to the application of identical mechanisms of relativisation. In this respect, they are similar to other Slavic languages. The four structural types of RCs – subject, object, indirect object and prepositional object – display similar syntactic patterning.

5.1 Subject RCs

In the Mac corpus, SRCs are more common than ORCs; they dominate over ORCs with the ratio 3:1 in literary prose and 4:1 in journalistic discourse. The analysis of the Bul literary corpus shows that SRCs prevail over ORCs with the ratio 2:1. In the journalistic corpus, SRCs are more frequent (6:1) than in literary prose. The widespread use of passive constructions may be responsible for the higher rate of SRCs in newspapers. Two syntactic patterns given in (8) and (9) are characteristic of SRCs: transitive and intransitive.

(8) SRC_{trans}: Head N + RO + V + (AccCl) + NP_(do/io)

(9) SRC_{intrans}: Head N + RO + V

The transitive pattern RO+Cl may contain an obligatory anaphoric accusative or dative clitic if the object NP is definite, as shown in (10).

(10) *Čovekot_j što/kojšto Ø_j brka edno kuče_i*
 man+the.M.Sg that/who.M.Sg chase.3Sg.Pres. one.Neut dog.Neut.Sg
/go_i brka kučeto. Mac
 3Sg.M.Acc.Cl chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man who is chasing a/the dog.’

Examples from the corpus illustrate this type:

(11) *Tie se poinakvi od želbite što ja*
 they are different from wish+the.Pl that she.Acc.Cl
dvižat glutnicata od negoviot vid. (J9) Mac
 move.3Pl.Pres pack+the.F.Sg of his kind
 ‘They are different from the desires that drive the pack of his kind.’

(12) *Slikarite koi gi prodavale svoite dela kaj* Mac
 painters+the.Pl who.Pl 3Pl.Acc.Cl sold.Pl./-Part own+the.Pl works at
nego ja podučivale da crta. (S34)
 him 3Sg.F.Acc.Cl taught.Pl./-Part Subj.Mark draw.3Sg.Pres
 ‘The painters who sold their paintings to him and taught her to paint.’

In Bul transitive SRCs, preverbal copy clitics of definite object NPs are optionally used in communicatively marked syntactic environments:

- (13) *Toj gi pribiraše (listata)... da si pridadat sila za dālgija pāt kām smārt, kojato gi očakvaše v kofite za bokluk.* (MS61) Bul
 he 3Pl.Acc.Cl collect.3Sg.Imperf leaves+the.Pl Subj.Mark Refl.Cl
 add.3Pl.PresPf strength for long journey to death.F which.F 3Pl.Acc.Cl
 await.3Sg.Imperf in bin+the.Pl for garbage

‘He collected the leaves... to give them strength for a long journey to death, that awaited them in the dustbins.’

It seems that a RC in the intransitive verb frame (9) is more tightly integrated in the main clause if the dependent verb is imperfective and if the RC is short. The degree of clause integration depends on the aspectual semantics of the modifying event and the immediate adjacency to the verb. A statal event becomes a defining property of the joint referent of the main event. As a result, a SRC with an imperfective verb is more tightly integrated within the head N than an ORC with identical configuration and acquires participial character. Below are examples from the corpus:

- (14) *Ķe bide izbran ušte mnogu pati vo godinite što doagaat.* (F3) Mac
 will.Mod.Cl. be.3Sg chosen.M.Part more many times in year+the.Pl
 that come.3Pl.Pres

‘He will be chosen many more times in the years to come.’

- (15) *Možebi taka se rağa onaa ljubov što dava.* (DS127) Mac
 perhaps so Refl.Cl bear.3Sg.Pres that love that give.3Sg.Pres
 ‘Perhaps that’s how the giving love is being born.’

- (16) *Pogledna nadolu i vidja ukazatelite, koito sočexa nadolu.* (MS50) Bul
 look.3Sg.Aor down and see.3Sg.Aor sign+the.Pl which.Pl
 point.3Pl.Imperf downwards

‘He looked down and saw the signs that pointed downwards.’

- (17) *I dobavix če ženite koito pušat, ne se polzvat s dobro ime.* (IP80) Bul
 and add.1Sg.Aor that women+the.Pl which.Pl smoke.3Pl.Pres Neg Refl.Cl
 use.3Pl.Pres with good name

‘And I added that the women who smoke have a bad reputation.’

5.2 Object RCs

The Mac ORC pattern includes an obligatory accusative clitic with a definite object NP and an optional subordinate subject NP; in Bul the clitic may be used for pragmatic

purposes. The corpus contains three Mac ROs: *koj*, (*kojšto*), *kogo*, *što* and three Bul ROs: *kojto*, *kogoto* and *deto* for object positions. The object syntactic pattern is given in (18); it is illustrated by example (19) with a marked empty position.

(18) ORC: Head N + RO + (AccCl) + V + (NP_{sub})

(19) *Čovekot_i što/kojšto_i (e_j) go_i brka Ø_i kučeto_j.* Mac
 man+the.M.Sg that/who.M.Sg 3Sg.M.Acc.Cl chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man that the dog is chasing.’

Examples from both corpora follow:

(20) *Vo golem broj slučai gi znaeme jadenjata*
 in great number cases 3Pl.Acc.Cl know.1Pl.Pres dish+the.Neut.Pl
što gi baraat. (F3) Mac
 that 3Pl.Acc.Cl ask.3Pl.Pres

‘In many cases we know the dishes that they order.’

(21) *Ogleda se dali njakoj ču glupostite*
 look around.3Sg.Aor Refl.Cl if someone hear.3Sg.Aor stupidity+the.F.Pl
koito izreče. (MS74) Bul
 which.Pl say.3Sg.Aor

‘He looked around to see whether anyone had heard the nonsense he said.’

(22) *Njakoj nepoznat...čovek, kogoto (Valeri) nenaviždaše*
 some unknown man who.Sg.Acc Valeri hate.3Sg.Imperf
ot cjalata si duša. (MS68) Bul
 from whole+the.F.Sg Refl.Cl.Dat soul.F.Sg

‘Some unknown man... whom he hated with all his heart.’

(23) *Vie kato im vidite kolite deto gi*
 you when 3Pl.Dat.Cl see.2Pl.Pres car+the.F.Pl that 3Pl.Acc.Cl
karat - povečeto bālgari karat vtora ruka koli.(coll.) Bul
 drive.3Pl.Pres most Bulgarians drive.3Pl.Pres second-hand cars

‘You should see the cars they are driving- and most Bulgarians drive second-hand cars.’

Depending on the choice of register, an indirect object NP can be relativized in three ways: (a) via dative declinable pronouns – in standard, mostly written language; (b) by a preposition + declinable pronoun in standard spoken language (cf. 24); (c) and, finally by *što* – in colloquial use (cf. 25). The dative RO *komu*, characteristic of the literary language and western dialects, in colloquial language is typically replaced by the prepositional RO *na koj*.

- (24) *Čovekot_i na koj/komu_i mu_i telefonira Ø_i včera*
 man+the.M.Sg to who.M.Sg/whom 3Sg.M.Dat.Cl call.2Sg.Aor yesterday
e tuka. Mac
 is here
 ‘The man who/whom you called yesterday is here.’

- (25) *Čovekot_i što mu_i telefonira Ø_i včera e tuka.* Mac
 man+the.M.Sg that 3Sg.M.Dat.Cl call.2Sg.Aor yesterday is here
 ‘The man that you called yesterday is here.’

In the Bul corpus prepositional ROs *na kogoto* and *na kojto* were found with dative subjects:

- (26) *Čovekāt na kogoto vsičko mu beše jasno*
 man+the.M.Sg to who.Sg.Acc everything 3Sg.M.Dat.Cl be.Sg.Imperf clear
za života i za smārtta. (MS174) Bul
 about life+the.M.Sg and about death+the.F.Sg
 ‘The man to whom everything about life and death was clear.’

The only examples with *deto* come from the internet, as (27):

- (27) *Tezi filmovi sa prax v očite na masovija zritel,*
 those movies be.Pl.Pres dust in eye+the.Pl. of mass+the.M.Sg viewer.M.Sg
deto mu daj lesnosmilaem action. (coll.) Bul
 that 3Sg.M.Dat.Cl give.Sg.Imp light action film.’
 ‘Those movies are like throwing dust into the eyes of the mass viewer who demands an easily digestible action film.’

6. Structural differences between SRCs and ORCs

The surface structure pattern of a SRC and an ORC is identical in transitive structures. Confusion between *što*-SRCs and *što*-ORCs in Mac may arise under two conditions: (a) when the subordinate object is definite and shares the same gender/number properties as the subordinate subject in a SRC; and (b) when the subordinate subject moves to postverbal position, leaving the subject position empty (e) in an ORC. The two structural patterns are represented in (28) nad (29):

(28) SRC: RO + Ø_{sub} + Cl_{ACC} + V + NP_{ob}

(29) ORC: RO + (e_{sub}) + Cl_{ACC} + V + Ø_{ob} + NP_{sub}

In example (30) the RC modifies the agent, the man who chased the dog, whereas in (31) the RC refers to the patient, the man chased by the dog.

(30) *Čovekot_j što Ø go_i brka kučeto_i.* Mac
 man+the.M.Sg that Sg.M.Acc.Cl chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man who is chasing the dog.’

(31) *Čovekot_j što (e_j) go_j brka Ø_j kučeto_i.* Mac
 man+the.M.Sg that 3Sg.M.Acc.Cl chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man that the dog is chasing.’

Transitive *što*-SRCs with a definite object have the same syntactic pattern as OSRs with *što*+Cl relativisation strategy. This is the reason why they can be interpreted as ORCs. In order to disambiguate (31) the subject NP must occupy the focus position (e_j), i.e., the slot after the RO, as in (32).⁸

(32) *Čovekot_i što kučeto_j go_i brka Ø_i.* Mac
 man+the.M.Sg that dog+the.Neut.Sg 3Sg.M.Acc.Cl chase.3Sg.Pres
 ‘The man that the dog is chasing.’

In Bulgarian, the same example may have two interpretations depending on their underlying syntactic patterns. In (33) the relative clause is SRC, while in (34) it is ORC. Although the clitic is optional in both examples, its omission seems to give the relative clause SRC interpretation.

(33) *Čovekāt_i deto Ø_i (go_i) goni kučeto_j.* Bul
 man+the.M.Sg that (3Sg.M.Acc.Cl) chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man who is chasing the dog.’

(34) *Čovekāt_i deto (go_j) goni Ø_i kučeto_j.* Bul
 man+the.M.Sg that (3Sg.M.Acc.Cl) chase.3Sg.Pres dog+the.Neut.Sg
 ‘The man that the dog is chasing.’

7. Possessive RCs

Sections 7 and 8 investigate the scope of structural divergence of Mac and Bul RCs with possessive and locative function. The analysis of the corpus shows that standard Mac does not allow the use of indeclinable *što* for possessive RCs (PossRC) and employs the declinable *čij* ‘whose’ or the prepositional RO *na koj(što)* ‘of who’. Bul possessive ROs are recruited from three sources: the prepositional ROs *na kogotolna kojtolot kojto* ‘of whom/of who’, the declinable *čijto* ‘whose’ and the indeclinable *deto* ‘that’. The syntactic pattern of PossRC is shown in (35) where NP_{poss} stands for the “possessor” NP. In the process of relativization, the “possessum” NP in square brackets (marked as Ø_{head}) is deleted due to its co-reference with the head N.

(35) PossRC: NP_{head} + RO + [NP_{poss} + Ø_{head}] + V

⁸ Rudin (1985:125) claims that Bul RCs resist co-occurrence with a filled focus position because Bul focused material in a RC is pragmatically odd.

The following examples illustrate the above pattern:

- (36) *Vo paragrafot... vleguva i oružjeto čij*
 in paragraph enter.3Sg.Pres and weapon+the.Neut.Sg whose.M.Sg
pogon e tetiva: lak i strela. (D1) Mac
 engine.M.Sg be.3Sg.Pres arch bow and arrow
 ‘The weapons that are based on arching: bow and arrow are included in the paragraph.’
- (37) *Vlizam v edna sgrada, čijšto zaden vxod*
 enter.1Sg.Pres in one building whose.M.Sg back entrance.M.Sg
mi e dobre izvesten. (BR162) Bul
 1Sg.Dat.Cl be.3Sg.Pres well known.Part.M.Sg
 ‘I enter a building whose back entrance I know very well.’
- (38) *Toj vidja ženata, na kojato broškata*
 he see.3Sg.Aor woman+the.F.Sg to who.F.Sg broach+the.F.Sg
beše prinadležala. (MS164) Bul
 be.3Sg.Imperf belonged.F.l-Part
 ‘He saw the woman who owned the broach.’
- (39) *Pier, taka se kazvaše čovekāt, na kogoto dopredi*
 Pierre thus Refl.Cl call.3Sg.Imperf man+the.M.Sg to who.Sg.Acc before
sedmica ne znaex imeto. (MS273) Bul
 week not know.1Sg.Imperf name+the.Neut.Sg
 ‘Pierre, that was the man, whose name I didn’t know a week ago.’
- (40) *Ili s izdisaštoto parata si kafe,* Bul
 or with exhuming+the.Neut.Sg steam Refl.Cl.Dat coffee.Neut.Sg
uxanieto ot koeto vārveše sled nego. (MS179)
 flavour+the.Neut.Sg of which+the.Neut.Sg follow.3Sg.Imperf after him
 ‘Or with the steaming coffee, the flavour of which followed him.’
- (41) *Tozi Simon, na kojto piše imeto tuk,*
 this Simon to who.M.Sg write.3Sg.Pres name+the.Neut.Sg here
e bil na sedemdeset godini. (MS189) Bul
 be.3Sg been.M.Sg.l-Partat on seventy years
 ‘This Simon, whose name is written here, was about seventy.’

The gapping strategy that relativizes a possessive NP in Bul is performed by the indeclinable RO *deto* followed by a “possessum” NP with a genitive clitic. The syntactic representation of this strategy given in (42) is illustrated by the only example in the Bul literary corpus (43):

(42) PossRC: NP_{head} + RO + [NP_{poss} + GenCl_i + Ø_i] + V

(43) *Imame edin šased, deto očite mu natičat,*
 have.1Pl.Pres one neighbour.M.Sg that eyes+the.Pl 3Sg.Dat.Cl bulge.3Pl.Pres
kato me gleda. (IP86) Bul
 when 1Sg.Acc.Cl look.3Sg.Pres

‘We have a neighbour whose eyes bulge when he looks at me.’

It seems that *deto*-relativisation in Bul is possible because of the presence of adnominal dative clitics. Bul possessive enclitics have no distributional constraints as they do in Mac where they can be used only with nouns that refer to family members in singular. Presumably, this constraint blocks the indeclinable RO+Cl relativisation. Some speakers when using possessive *što*-relativisation (in colloquial Mac), resort to the retention of the anaphoric pronoun after the preposition *od* ‘from’ as the following example illustrates:

(44) *Stignavme do rekata što pievme voda od nea. (coll) Mac*
 arrive.1Pl.Aor to river+the.F.Sg that drink.1Pl.Imperf water from she.Acc
 ‘We arrived at the river we had drunk water from.’

8. Prepositional RCs

In both languages the relativisation of oblique objects is commonly performed by the relative pronoun strategy. PrepROs consist of a preposition and a declinable pronoun *koj/kojto*. The preposition encodes the case role of the head PP (goal, location, time, manner, instrument, etc) and governs the declinable pronoun (*kogo/kogoto*, the accusative form of *kojto*). In colloquial Mac, the nominative *koj(što)* is preferred instead of the accusative *kogo*. The relativisation of prepositional objects via *kojšto/kojto* involves two operations: (a) the NP complement of the preposition in the PP is replaced by a declinable relative pronoun which results in a compound RO (PrepRO) and (b) the prepositional RO (PrepRO) is moved to clause-initial position.

PrepROs in Mac and Bul relativize a similar range of oblique positions: indirect object, possessive phrase and modifying adverbials with locative, temporal and manner functions. The gapping indeclinable strategy in Bul has a limited range of application in relativising the object of preposition. Only several examples with *deto* and a deleted sociative preposition *s* ‘with’ or the topical preposition *za* ‘about’ were encountered in the corpus. In (46) the relative clause is non-restrictive:

(45) *Vi predupredja za nešto, deto ne znaete*
 2Pl.Dat.Cl warn.1Sg.Aor about something that Neg know.2Pl.Pres
dokāde mu stigaat korenite. (KD57) Bul

to-where 3Sg.Dat.Cl reach.3Pl.Pres root+the.M.Pl

‘I warned you about something that you don’t know how deep its roots are.’

- (46) *Tova e Emil, deto bjagaxme zaedno.* (BR45) Bul
 this be.3Sg.Pres Emil that run.1Pl.Imperf. together
 ‘This is Emil, who we used to run away with.’

Gapping, i.e., the use of indeclinable RO without a clitic (RO-Cl), is often combined with the pronoun retention strategy. The retention of a “long” pronoun enables relativization of a PP, as the following Bul examples illustrate:

- (47) *Oblekox staroto ti sako, složix onija kasket, deto xodiš s nego na ribolov...* (IP33) Bul
 dress.1Sg.Aor old+the.N.Sg 2Sg.Dat.Cl jacket.N.Sg put.1Sg.Aor that.M.Sg
 hat.M.Sg that go.2Sg.Pres with he.Acc to fishing
 ‘I put on you old jacket, put on the hat that you go fishing with.’

- (48) *Decata, deto igraexme s tjax ...* (from Rudin 1986:139) Bul
 children+the.N.Pl that play.1Pl.Imperf with they.Acc
 ‘The children that we played with’

Some examples in the corpus are ambiguous between the locative *deto* and the relative *deto*, due to deletion of the preposition *vo* ‘in’ as in (49):

- (49) *Mojat Pariz njama ništo obšto s tozi pātevoditel, det figurira Ajfelovata kula.* (BR57) Bul
 my+the.M.Sg Paris have.3Sg.Pres nothing common with this guide.M.Sg
 that represent.3Sg.Pres Eiffel+the.F.Sg Tower.F.Sg
 ‘My Paris has nothing to do with this guide where the Eiffel Tower is shown.’

9. Types of elements eligible for relativisation

The analysis of both corpora has shown that the range of the indeclinable relativisation in Mac is smaller than in Bul. The invariant *što*-relativisation strategy in Mac is restricted mainly to subject and direct object positions, whereas Bul *deto* allows relativisation from almost all syntactic positions. The syntactic properties of the indeclinable ROs may be established using Keenan and Comrie’s hypothesis (1977) about the preferred order of the relativisation positions, also known as Noun Phrase Accessibility Hierarchy - AH. For this purpose Mac and Bul ROs are compared following Joseph’s (1982) analysis of Greek RCs. The contrasted relativisation positions in AH are represented in table 1:

Table 1

RO	SUBJ	DO	IO	POSS	PrepO
----	------	----	----	------	-------

što/deto	+	+	+	??/+	??/+
kojto/koj(što)	+	+	+	+	+

The above table and the examples in section 7 and 8 show that *što*-relativisation has a limited application in Mac. Bul *deto* may relativize from more syntactic positions including the possessive and the prepositional object positions. Yet, most accessible for relativisation are the subject and the object positions. As far as the use of clitic strategy is concerned, the situation in Mac is quite clear: it is the type of RC that determines the choice between the strategy with a clitic (+Cl) or without the clitic (-Cl). Mac subject *što*-clauses always use RO-Cl strategy, while object *što*-clauses employ RO+Cl strategy. Having a communicative role, clitics are optional in oblique positions in Bul RCs. Below, several examples from the corpus illustrate the *deto*-relativisation: subject in (50), object in (51), possessive in (52), and prepositional object in (53).

- (50) *Ima bambini od našija grad, deto redovno*
 have.3Sg.Pres dude.Pl from our+the.M.Sg town.M.Sg that regularly
si prekarvat ljatoto na more. (IP74) Bul
 Refl.Cl. Dat spend.3Pl.Pres summer+the.Neut.Sg at sea
 ‘There are dudes in our town who regularly spend the summer at the seaside.’
- (51) *Bibliotečkata, deto pobiraše Jugo i Stendala,*
 library+the.F.Sg that contain.3Sg.Imperf Hugo and Stendhal
včera trjabvaše da se rasširi. (KD28) Bul
 yesterday need.3Sg.Imperf Subj.Mark Refl.Cl expand.3Sg.Pres.Pf
 ‘The library, that had Hugo and Stendhal was supposed to be expanded yesterday.’
- (52) *Iskam da kaza na decata deto*
 want.1Sg.Pres Subj.Mark tell.1Sg.Pres.Pf to children+the.Neut.Pl that
pārvata kniga deto sa xvanali se e
 first+the.F.Sg book.F.Sg that be.3Pl held.3Pl.l-Part Refl.Cl be.3Sg
kazvala Xari Poter. (coll) Bul
 named.F.Sg.l-Part Harry Potter
 ‘I want to tell the children whose first book that they came across was Harry Potter.’
- (53) *V Dobrudža njamaše kāšta, deto krakāt*
 in Dobrudža not-have.3Sg.Imperf house.F.Sg that foot+the.M.Sg
mu da ne e stāpval. (IP13) Bul
 3Sg.Dat.Cl Subj.Mark Neg be.3Sg stepped.M.Sg.l-Part
 ‘There was no house in Dobrudža where he hasn’t set his foot.’

The comparison of the relativisation strategies in two languages yields the following results:

(a) relativisation with a variable RO is possible from all positions in both languages; (b) Mac SRCs always use RO-Cl strategy, while ORCs employ the RO+Cl strategy; (c) the indeclinable *deto* in Bul is limited, being restricted to colloquial use but has a wider range of syntactic applications.

10. Declinable vs. indeclinable ROs

Mac declinable *koj* is commonly used in journalistic prose in presentative (54) and existential constructions (55). In the former, it explains the unspecified head N.

(54) *Toa e složena dejnost koja bara poinakov*
 this is complex.F.Sg activity.F.Sg which.F.Sg demand.3Sg.Pres different
pristap. (D) Mac
 approach

‘This is a complex activity that demands a different approach.’

(55) *Vo svetot postojat tri firmi koji go*
 in world+the.M.Sg exist.3Pl.Pres three firm.Pl which.Pl 3Sg.Acc.Cl
proizveduvaat ovoj lek. (V) Mac
 produce.3Pl.Pres this.M.Sg medicine.M.Sg

‘There are three firms in the world that produce this medicine.’

The analysis of the examples shows that the choice of register influences the distribution of *koj*-clauses whose number in journalistic prose is almost twice higher than in literary prose. Structural criteria are also important in that *koj*-SRCs dominate over *koj*-ORCs in both genres with ratio 5:1. The referent of the head N can be both personal and non-personal, though personal referents are more common.

In Bul prose, the declinable *kojto* overwhelmingly prevails over the indeclinable *deto*. The latter is used in stylistically marked discourse to convey familiarity. In Mac the functional border-line between the two Ros is blurred. Yet, the analysis of the *koj*-examples from the corpus allows to establish some tendencies. The choice of RO appears to depend on the length and the structural type of the RC. *Koj* is preferred with SRCs modifying subject heads, except when the RC has a participial character (see examples 14 and 15). This is evident in simple “short” subject NPs where *što*-RCs function as active participles (56).

(56) *Čovekot što doaga e moj sosed.* Mac
 man+the.M.Sg that come.3Sg.Pres be.3Sg my.M.Sg neighbour
 ‘The man that is coming is my neighbour.’

(57) *?Čovekot što doaga po patekata e moj sosed.* Mac
 man+the.M.Sg that come.3Sg.Pres on path+the.F.Sg be.3Sg my.M.Sg neighbour
 ‘The man that is coming down the path is my neighbour.’

(58) *??Čovekot koj doaga e moj sosed.* Mac

man+the.M.Sg who.M.Sg come.3Sg.Pres be.3Sg my.M.Sg neighbour
 ‘The man who is coming is my neighbour.’

- (59) *Čovekot koj doađa po patekata e*
 man+the.M.Sg who.M.Sg come.3Sg.Pres on path+the.F.Sg be.3Sg
moj sosed. Mac
 my.M.Sg neighbour
 ‘The man who is coming down the path is my neighbour.’

It seems that clause integration is higher with *što* than with *koj*. SRCs that modify subject heads tend to co-occur with a restrictive *koj*. As (60) illustrates, *koj* tends to give an appositive reading to the RC that modifies the object head N in the SRC:

- (60) *Go znam čovekot što/?koj radi tuka.* Mac
 3Sg.Acc.Cl know.1Sg.Pres man+the.M.Sg that/who.M.Sg work3Sg.Pres here
 ‘I know the man who/that works here.’

Although Mac prescribes the use of *što*, today *koj* (and *kojšto*) have encroached into the functional zone of *što*. In current language usage, younger generations prefer *koj* to *što*. The data in the corpus supports the observation that this phenomenon is an ongoing tendency. It is not accidental that Janeski (J) in his novel from the early 70-ies avoids *koj*, while Smilevski (S), who wrote his book almost forty years later, does the opposite. Korubin (1990:268) notes that the first and second generation of writers after the codification of the language, predominantly used *što* even in cases of apposition illustrating this by example in (61). In contrast, Smilevski uses *koj* even when *što* is more suitable than *koj*, as it is in (62):

- (61) *Veljko Vlahovic, što ja predvodi delegacijata...* Mac
 Veljko Vlahovic, that 3Sg.Acc.Cl lead.3Sg.Pres delegation+the.F.Sg
 ‘Veljko Vlahovic, who is leading the delegation...’

- (62) *Na potkrovniot kat... tatko čuvaše neki od vinata*
 on top+the.M.Sg floor.M.Sg father keep.3Sg.Imperf some of wine+the.Pl
koi gi prodavaše vo negovata prodavnica. (S22) Mac
 which.Pl 3Pl.Acc.Cl sell. 3Sg.Imperf in his+the.F.Sg shop.F.Sg
 ‘On the top floor, father kept some of the wines that he sold in his shop.’

The spread of *koj* into the functional zone of the *relativum generale* *što* in spoken and written standard Mac reflects a divergent process that departs from a common Balkan pattern. One can only hypothesize about the reasons for greater popularity of *koj* and *kojšto* at the expense of *što*, especially with younger generations. The expansion of the functional zone of *koj* may be actually attributed to two opposing tendencies: (a) a tendency for simplification: speakers use only one relativizer for both restrictive and non-restrictive modification; and (b) a tendency for disambiguation: *što* occurs in relative clauses that have a clear restrictive sense; all other unclear cases are solved with *kojšto*.

If Mac tends to restrict the use of the indeclinable RO at the expense of the declinable relativizer, Bul makes a stylistic differentiation reserving *deto* for the colloquial use only, as examples (63) and (64) illustrate.

- (63) *Imaše xora, koito ne viždaxa ništo pred*
 have.3Pl.Imperf people who.Pl Neg see.3Pl.Imperf nothing before
sebe si. (IP100) Bul
 own Acc. Refl.Cl.Dat
 ‘There were people who saw nothing in front of them.’

- (64) *Mu kazvali, če ima xora, deto bjagat ot*
 3Sg.Dat.Cl told.Pl.l-Part that have.3Sg people that run.3Pl.Pres from
dobroto kato zajci ot kopoj. (IP103) Bul
 good+the.Neut.Sg like rabbit.Pl from hound.M.Sg
 ‘He was told that there are people who run away from good like rabbits from a hound.’

Bul *deto* can also be used in appositive RCs:

- (65) *Raspoznax glasa na bradatija Stefan, deto*
 recognize.1Sg.Aor voice+the.M.Sg of bearded+the.M.Sg Stefan that
minalija sezon igra Makbet. (KD95) Bul
 previous+the.M.Sg season play.3Sg.Aor Macbeth
 ‘I recognized the voice of bearded Stefan, who played Macbeth in the previous season.’

11. Quantitative analysis of the distribution of *koj* vs. *što*

11.1 Data from journalistic sources

The summary quantitative analysis presented in table 2 shows that *koj* in Mac is used marginally with a restrictive function; *što* occurs more than *koj*, but the distribution varies depending on type of source. The two major newspapers are indicative of contemporary usage: the more conservative *Dnevnik* follows the prescribed norm to use *što*, whereas the popular *Vest* reflects the colloquial usage tendency to overuse *koj*. The two journals also display a similar pattern probably because the articles in them were authored by younger journalists. In addition, the analyzed text from the *Kapital* contained numerous existential constructions that usually co-occur with *koj*.

Table 2: The distribution of *koj* and *što* in Mac journalistic prose

source	KOJ	STO	KOJSTO
Dnevnik	10	43	
Vest	32	27	

Forum	24	44	3
<u>Kapital</u>	<u>22</u>	<u>10</u>	<u>1</u>
Total (200)	88	114	4

The analysis of the examples in the journalistic prose corpus shows that *koj* tends to be used (a) with SRCs including existential constructions and nominal predicates, (b) with human referents, and (c) with indefinite referents. Below each of the three factors that influences the distribution of *koj* is elaborated in more detail:

(a) favourable syntactic environment

SRCs appear to be a favourable syntactic environment for *koj*-clauses. Thus, only seven ORCs were found in all *koj*-clauses in the journalistic corpus (cf. 66). Other favourable syntactic environment for *koj*-clauses are existential constructions (cf. 67) and nominal predicates (cf. 68).

- (66) *Imame stranski klienti koi baraat hrana što*
 have.1Pl.Pres foreign.Pl customer.Pl who.Pl want.3Pl.Pres food.F that
ja nema na našeto meni. (F3) Mac
 3Sg.Acc.Cl not-have.3Sg.Pres on our+the.Neut.Sg menu.Neut.Sg

‘We have foreign customers who want food that is not on the menu.’

- (67) *Čovekot bez cel e čovek koj im Mac*
 man the.M.Sg without goal be.3Sg.Pres man who.M.Sg 3Pl.Dat.Cl.
se prepušta na slučajnite čuvstva ili neidentifikovani nagoni. (A)
 Refl.Cl submit.3Sg.Pres to accidental feelings or unidentified instincts

‘The man without a goal is a man who submits himself to accidental feelings or unidentified instincts.’

(b) human referent of head N in *koj*-clauses

The head N in *koj*-clauses tends to have a human referent. The ratio of human vs. non-human referents in *koj*-clauses is 4:1. But there are examples with non-human referents such as (68).

- (68) *Poslednata godina vekje čuvstvuvaše želba koja*
 last+the.F.Sg year.F.Sg already feel.3Sg.Imperf desire which.F.Sg
se poteško se kontrolira. (F3) Mac
 totally difficult.Comp Refl.Cl control.3Sg.Pres

‘During the last year he already felt a desire that was becoming more difficult to control.’

(c) indefinite head N in *koj*-clauses

The head N in *koj*-clauses is very often indefinite as in (69). In such cases the head denotes a non-referring or unspecified object, sometimes modified by indefinite quantifiers

the *nekoj* ‘someone’, *mnogumina* ‘a lot’ (70). However, there are some examples such as (71), where the *koj*-clause modifies a definite head.

(69) *Toj go minuva svojot život barajki vrednosti*
 he 3Sg.Acc.Cl. spend.3Sg.Pres own+the.M.Sg life.M.Sg searching values
koi nikogaš nema da gi najde. (A) Mac
 which.Pl never Neg Subj.Mark 3Pl.Acc.Cl find.3Sg.Pres.Pf
 ‘He spends his life looking for values that he will never find’

(70) *Na Kongresot na nekoi kolegi koi se lutea*
 at congress+the.M.Sg to some colleagues who.Pl Refl.Cl upset.3Pl.Imperf
im rekov deka jas ne glasam za G. (V1) Mac
 3Pl.Dat.Cl tell.1Sg.Aor that I Neg vote.1Sg.Pres for G.
 ‘At the Congress I told some colleagues who were upset that I wouldn’t vote for G.’

(71) *Na koleškata koja gi podgotvuvaše dogovorite*
 to colleague+the.F.Sg who.F 3Pl.Acc.Cl prepare.3Sg.Imperf contract+the.Pl
i rekov da go prezeme
 3Sg.F.Dat.Cl tell.1Sg.Aor Subj.Mark 3Sg.M.Acc.Cl take-over.3Sg.Pres.Pf
ovoj. (K) Mac
 this.M.Sg
 ‘To the colleague who was preparing the contracts I told to take over this one.’

Some *koj*-clauses are ambiguous between restrictive and appositive reading:

(72) *Otsustvuvala samo učenici od sosednite sela*
 missed.Pl.l-Part only pupil.Pl from neighbouring+the.Pl village.Pl
koi imale problem so prevozot. (V2) Mac
 who.Pl had.Pl.l-Part problem with transportation+the.M.Sg
 ‘Only the pupils from the neighbouring villages were absent, who had problems with transportation.’

11.2 Data from literary sources

The figures in parentheses in table 3 and table 4 denote the common number of subject and object relatives in literary corpora in the two languages. In Mac literary sources the ratio of *što* vs. *koj* is 2:1. However, the distribution varies to a significant degree between recent and older novels. In Smilevski (S), *koj* prevails over *što* with the ratio 5:1, whereas in earlier analyzed works from the corpus, the opposite is true. Thus, in Janevski (J) *koj* is marginally used, and in Solev (DS) *što* dominates *koj* with the ratio 5:1.

Table 3: Macedonian literary sources

source		<i>koj</i>	<i>što</i>
J		3(2+1)	55(35+18)
DS		10(9+1)	59(47+12)
S		42(35+7)	
Total	200	63	137

The distribution of the Mac declinable relativizer *kojto* and its Bul counterpart *deto* in the Bul corpus strongly depends on the register. The pressure of the norm to avoid *deto* in literary discourse has led to its marginalisation. Hence it is more often found in older sources; if the ratio of Bul *kojto*: *deto* in the books written in the seventies (Bogomil Rajnov – BR, Ivajlo Petrov – IP) is 6:1 and 7:1 respectively, in a recently written book (Kristin Dimitrova – KD) *deto* is marginally represented (with 15:1). It is worth noting that *deto* in all four sources is used in dialogues, which supports the assertion about its colloquial character.

Table 4: Bulgarian literary sources

source	<i>kojto</i>	<i>deto</i>
IP	119(95+34)	17
KD	90(59+31)	6
Total (232)	209(154+65)	23

The situation is different in the corpus compiled from Bulgarian newspapers. There is no occurrence of *deto* in 200 examples in that corpus. The absence of *deto* in written language is in sharp contrast with its frequent use in colloquial discourse in various Bulgarian internet chat rooms. The following examples (73-75) show that *deto* co-occurs with human and non-human referents, in subject and object RCs.

(73) *Teško im e na mladite deto štiaxa da*
 difficult 3Pl.Dat.Cl be.3Sg.Pres to young+the.Pl that will.Mod.Cl Subj.Mark
kupuvat i da se drogirat. (coll.) Bul
 buy.3Pl.Pres and Subj.Mark Refl.Cl drug.3Pl.Pres

‘Too bad for those young people who wanted buy and use drugs.’

(74) *Firmata deto mi izdejstva zelenata karta, iska*
 firm+the.Sg.F that 1Sg.Dat.Cl issue.3Sg.Aor green card want.3Sg.Pres
da me vzeme za tezi 39500 Eur. (coll) Bul
 Subj.Mark 1Sg.Acc.Cl employ.3Sg.Pres.Pf for those 39.500 euros

‘The firm that helped me get the green card wants to employ me for 35.000 euros.’

- (75) *Kažete edna banka, deto njama da mi otkažat!* (coll.)Bul
 tell.2Pl.Imp one bank that Neg Subj.Mark 1Sg.Dat.Sg reject.3Sg.Pres.Pf
 ‘Tell me the name of a bank that won’t turn me down!’

11.3 *Distribution of što-clauses*

The quantitative analysis of *što*-clauses has made it possible to reach three important generalisations about the distribution of *što*-clauses: (a) the preferred structural type where *što*-clauses frequently occur are SRCs, but *što*-ORCs can be found in the corpus as well. Subject *što*-clauses are illustrated in section 5.1, examples (11-13), while object *što*-clauses in examples (20-23); (b) *što*-clauses tend to occur with inanimate referents with the ratio 9:1 as shown in (76), though there are examples with human referents (cf. 77); (c) *što*-clauses tend to co-occur with definite referents, as in (78).

- (76) *Nie postojano se zanimavame so raboti što ni*
 we constantly Refl.Cl deal.1Pl.Pres with thing.Pl that 1Pl.Dat.Cl

pravat sopki. (K) Mac
 make.3Pl.Pres obstacles

‘We are constantly occupied with things that create obstacles for us.’

- (77) *Zemjodelcite što dobivaat takvo zemjište smetaat*
 farmer+the.Pl that get.3Pl.Pres such land think.3Pl.Pres

deka se izigrani. (D2) Mac
 that be.3Pl.Pres cheated.Pl.Part.

‘The farmers who get such land think that they have been cheated.’

- (78) *Taka veli lektorot što pred nekolku godini*
 thus say.3Sg.Pres instructor+the.M.Sg that before some year.Pl

se vratil od svojata misija. (A) Mac
 Refl.Cl returned.M.I-Part from own+the.F.Sg assignment.F.Sg

‘These are the words of a language instructor who had come back from his assignment a few years ago.’

Existential constructions and nominal predicates are not favoured by *što*-clauses, though several examples representing these two types can be found in the corpus:

- (79) *Vo policijat ima lica što ne zaslužuvaat da*
 in police exist.3Sg.Pres person.Pl that Neg deserve.3Pl.Pres Subj.Mark

ja nosat uniformata. (D3) Mac
 3Sg.Acc.Cl wear.3Pl.Pres uniform+the.F.Sg

‘There are policemen who do not deserve to wear the uniform.’

- (80) *Toa e film što zboruva za*

this be.3Sg.Pres movie.M.Sg that say.3Sg.Pres about
baranjeto na radost. (D1) Mac
 search+the.Neut.Sg of joy

‘This is a movie that speaks about the search for joy.’

12. RCs with pronominal heads

Pronominal relative clauses in Mac can be divided into two groups depending on the type of pronoun that heads the RC. In the first group the heads are formalized as definite demonstrative pronouns (distal *onoj*, neutral *toj* and proximal *ovoj*), while the second group involves the indefinite pronouns (*nekoj* ‘somebody’, *sekoj* ‘everybody’, *se* ‘everything’, *sie* ‘all’) as heads. In both languages pronominal relatives ground the unspecified referent of the pronoun in the discourse domain of the speaker cataphotically. These RCs acquire complement-like properties at the expense of modifying ones, since they are tightly integrated into the main clause.

(81) *Nikoj nigde ne gi spomenuvaše tie što*
 no one nowhere Neg 3Pl.Acc.Cl mention.3Sg.Imperf those that
živeeja vo seloto. (J144) Mac
 live.3Pl.Imperf in village+the.Neut.Sg

‘Those who lived in the country weren’t mentioned anywhere by anyone.’

(82) *Tezi, koito sega minavaxa po koridora, bjaxa*
 those which.Pl now pass.3Pl.Imperf along corridor+the.M.Sg be.3Pl.Imperf
kračkite na aktrisata. (MS151) Bul
 foot+the.Pl of actress+the.F.Sg

‘What was now walking down the corridor, were the feet of the actress.’

As mentioned above, the first type of pronominal RCs are headed by definite pronouns. Although pronominal heads co-occur both with indeclinable and declinable ROs, there are some preferred combinations. The analysis of the Mac corpus and the examples on the internet led to the following conclusions: (a) the occurrences of pronominal ROs with indeclinable *što* by far outnumber those with the declinable ROs *koj* and *kojšto*; (b) the singular declinable relativizer *koj*, *koja* and *koe* (though not the plural relativizer *koi*) co-occur with distal pronouns more readily than with proximate ones; (c) the use of *kojšto* is marginal in comparison with *koj*.

Bulgarian, on the other hand, prefers the indeclinable RO with all pronominal heads except the deictic *tova* ‘that’ (83). The relatively frequent co-occurrence of *tova* and *deto* probably has to do with the non-human reference of *tova*. But in literary styles the declinable RO is preferred (84):

- (83) *Mda, mnogo xubavo tova, deto si go*
 yeah very nice that.Neut.Sg that be.2Sg.Pres 3Sg.Acc.Cl
napisal. (coll) Bul
 written.M.Sg.l-Part
 ‘Yeah, what you have written is very nice.’
- (84) *Taka če ne sme tačno tova, koeto može*
 so that Neg be.1Pl.Pres exactly that.Neut.Sg which.Neut.Sg can.3Sg.Pres
da se nareče “čisti”. (Dn) Bul
 SubjMark Refl.Cl call.3Sg.Pres.Pf clean
 ‘So, we are not exactly what you might call “pure”.’

The second type of pronominal RCs involves clauses headed by indefinite pronouns *nekoj* ‘someone’, *nešto* ‘something’, *sekoj* ‘every’, *site* ‘all’, *se* ‘everything’, *nikoj* ‘no one’, *ništo* ‘nothing’ and *eden* ‘one/a’. The examples on the web have confirmed that these pronouns tend to associate with the indeclinable *što*, though the personal *nikoj* and *sekoj* might constitute an exception. The co-occurrence of the indefinite pronouns with the relativizer *koj* may be due to the pronounced “humanness” in the semantics of *nikoj* and *sekoj*.

It is important to underline that indefinite pronouns *sekoj*, *site*, *nikoj* and *ništo* function as quantifiers of the missing but implied nominal heads *man* and *thing*. The RC ascribes some property to the number of participants determined by the quantifiers ‘all’ or ‘no-one’. The indefinite *nekoj* and *nešto*, on the other hand, behave differently. They ascribe the property coded by the RC to unspecified pronominal heads. With *nekoj/nešto* as matrix head N, the RC grounds the new participant in the discourse cataphorically.

- (85) *Dojde nekoj što te znae.* Mac
 come.3Sg.Aor somebody that 2Sg.Acc.Cl know.3Sg.Pres
 ‘Somebody who knows you has come.’
- (86) *Nekoj što znae kompjuteri dobro zarabotuva.* Mac
 someone that know.3Sg.Pres computer.Pl well earn.3Sg.Pres
 ‘Someone who is familiar with computers earns well.’
- (87) *Mi kaža nešto što me začudi.* Mac
 1Sg.Dat.Cl tell.3Sg.Aor something that 1Sg.Acc.Cl surprise.3Sg.Aor
 ‘He told me something that surprised me.’

Structurally, pronominal RCs belong to relative clauses, because they share a joint constituent with the main clause and immediately follow the head. However, the pronominal RCs specify, rather than modify the head – they explain the content of the demonstrative pronoun via another proposition coded by the RC. Since they identify what or who the unspecified referent of the pronoun is, the behaviour of pronominal RCs resembles complement clauses. Therefore, in cases of cataphorical grounding the link between the pronominal RC and the main clause seems to be tighter.

13. Linearization in RCs

Two syntactic patterns of surface linearization of RCs can be distinguished in both languages, which differ with respect to the filling of a focus position to the immediate right of the RO. Pattern (A) represents unmarked linearization, while pattern (B) is pragmatically marked, because the focus position is occupied by a moved constituent. As the examples below show, the underlined focused constituent is an adverbial in (90, 91), subject NP in (92), or a whole clause in (93, 94).

(88) pattern A: RO + (Cl) + VP + NP(subj/obj) + AdvP

(89) pattern B: RO + AdvP/ NP(subj/obj)/S + Cl + VP

The following examples illustrate marked linearization of RCs, i.e., pattern B:

(90) *Luġeto* *vo minatoto* *baraa* *nekoi* *vistini* *što*
 people+the.Neut.Pl in past+the.M.Sg search.3Pl.Imperf some thruth.Pl that
i den denes *im* *se* *nepoznati.* (J11) Mac
 and today 3Pl.Dat.Cl be.3Pl.Pres unknown

‘People in the past were searching for some truths tha have remained unknown to them until now.’

(91) *S* *tezi* *pari* *možeše* *da* *si*
 with this.Pl money.Pl can.3Sg.Imperf Subj.Mark Refl.Cl.Dat
kupi *učebnici* *koito* *na staro dosega* *mu*
 buy.3Sg.Pres.Perf textbooks which.Pl second-hand until-now 3Sg.Dat.Cl
dostavjaše *Kostaki.* (MS133) Bul
 provide.3Sg.Imperf Kostaki

‘With this money he could buy textbooks, which until now Kostaki provided to him second-hand.’

(92) *No Burok* *ne* *zema* *ni* *zalak* *leb,* *ni što* *da*
 but Burok Neg take.3Sg.Pres nor morcel bread nor what Subj.Mark
e *što* *dobronamernicite* *mu* *nudat.* (J100) Mac
 be.3Sg.Pres that well wisher.Pl 3Sg.Dat.Cl offer.3Sg.Pres

‘But Burok does not take even a morcel of bread or anything else that the well-wishers offer him.’

(93) *I* *toj* *ja* *položi* *dlankata* *vrz* *čelo* *kako* *čovek*
 and he 3Sg.Acc.Cl put.3Sg.Aor hand+the.F.Sg on forehead like man
što *iako* *znael* *i* *rešil* *mnogu tajni* *na*
 that although known.3Sg.M./Part and solved.3Sg.M./Part many secrets of

referents. Pragmatic factors such as the need for a universal RO, both restrictive and non-restrictive, are responsible for the opposing tendency to replace *što* with *koj* or *kojšto*. As noted, the declinable relativizer *koj* is used in apposition, which is in line with the hypothesis of its increasing anaphoric function. As the Mac corpus indicates, this synthetic strategy is especially widespread among younger speakers, who tend to expand the synthetic *koj*-strategy at the expense of the analytic in the standard language.

The second type of accommodation of RCs to the head N within the matrix clause involves analytic clause linkage combined with clitic pronoun retention (indeclinable ROs+Cl). In both languages, unstressed clitic pronouns gravitate toward the second sentential position (cf. Comrie 1989:150). This can give rise to a mixed type between pronoun retention (full personal pronoun) and the relative pronoun type: a hybrid RO+Cl strategy common in both languages, especially in Mac.

The scope of indeclinable relativisation in Mac is smaller than in Bul. While in Mac the indeclinable relativization is restricted to subject and direct object positions, in Bul it is allowed from almost any syntactic positions. The wider functional load of *deto* may be ascribed to the unrestricted adnominal use of possessive clitics in Bul.

The two languages also differ with respect to stylistic restrictions on the application of the analytic strategy: Bul restricts the analytic strategy to the colloquial language, while Mac prescribes it. Although the synthetic strategy in Mac is encroaching into the functional zone of the analytic strategy, there are cases where the latter remains the only choice. Thus, the indeclinable strategy is used in spoken language, where it emphasizes the restrictive meaning of the embedded RC.⁹ Namely, the complement-like *što* is used in clear cases of restrictive clause linkage, while *kojšto* and *koj* alternate in all other cases.

The third type of relative clause linkage, marginally found in both languages, was referred to as “gapping” in the previous discussion. Gap relativisation involves cases when the indeclinable RO stands without its clitic part (RO-Cl), but is often followed by the full pronoun copy. Pronoun retention (other than the resumptive one, which is clitic) is introduced in order to ease the relativisation of the most inaccessible, i.e., oblique positions, such as prepositional objects. Gapping is quite uncommon – in both languages it is found only in colloquial usage, though more often so in Bul. The few possessive RCs and PrepRCs in the corpus (cf. 43-49) illustrate relativisation from oblique positions.

The comparative corpus-based analysis has shown that Mac and Bul RCs are similar in two respects: (a) they share similar syntactic patterning of the four structural types of RCs – subject, object, indirect object and prepositional object; and (b) they possess the same inventory of relativisation strategies. However, the uneven distribution of the three types of ROs in these two languages indicates that the degree of application of these strategies is different. The functional zone of each strategy is determined by a combination of the structural factors discussed above and pragmatic factors such as register and adherence to the standard norm.

Abbreviations

relative clause	RC	NP head, antecedent of RC	head N
relative operator	RO	declinable RO+Cl	<i>koj</i> +clitic

⁹ Koneski (1987:540) thinks that *kojšto* is stylistically more inadequate than *koj*, which is common in the literary language.

subject relative clause	SRC	indeclinable RO+Cl	što+clitic
object relative clause	ORC		
prepositional rel. clause	PrepRC		
possessive relative clause	PossRC		

1/2/3 = 1st/2nd/3rd person; Acc = accusative (case); Aor= aorist (tense); Cl = clitic; Comp = Comparative; Dat = dative (case); F = feminine; Imp = imperative; Imperf = imperfect (tense); M = masculine; Mod = modal; Neg = negation; Neut = neuter; Part = participle; Pf = perfective (aspect); Perf = perfect (tense); Pl = plural; Pres = present (tense, imperfective aspect); Pres.Pf = present (tense, perfective aspect); Refl.Cl = reflexive clitic; Sg = singular; Subj.Mark = subjunctive marker.

References

- Browne, Wayles. 1981/1986. *Relativna rečenica u hrvatskom ili srpskom jeziku u poređenju s engleskom situacijom*. Doctoral dissertation, University of Zagreb, 1981. Published as Browne, Wayles. *Relative Clauses in Serbo-Croatian; The Yugoslav Serbo-Croatian - English Contrastive Project; New Studies 4*. [ed. Rudolf Filipović] Zagreb: Institute of Linguistics, 1986. [English translation by the author.]
- Bužarovska, Eleni. 2004. "The distribution of subjunctive relatives in Balkan languages". In Olga Mišeska Tomić, ed. *Balkan Syntax and Semantics*. Amsterdam: John Benjamins. 377-404.
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.
- Givón, Talmy. 1990. *Syntax. A Functional-Typological Introduction*. Amsterdam: John Benjamins.
- Joseph, Brian D. 1982. "Relativisation in Modern Greek. Another look at the accessibility hierarchy constraints". *Lingua* 60. 1-24.
- Keenan, Edward L., and Bernard Comrie. 1977. "Noun Phrase accessibility and universal grammar". *Linguistic Inquiry* 8/1. 63-99.
- Koneski, Blaže. 1987. *Gramatika na makedonskiot jazik*. [Grammar of Macedonian Language] Skopje: Kultura.
- Korubin, Blagoja. 1990. *Na makedonsko gramatički temi*. [On issues of Macedonian grammar] Skopje: Institut za makedonski jazik "Krstev Petkov Misirkov".
- Rudin, Catherine. 1985. *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*. Columbus: Slavica Publishers.
- Penčev, Jordan. 1984. *Stroez na bălgarskoto izrečenie*. [The Structure of the Bulgarian Sentence] Sofija: Nauka i Izkustvo.
- Pašov, Petăr. 1989. *Praktičeska bălgarska gramatika*. [Practical Bulgarian Grammar] Sofija: Hermes.
- Topolińska, Zuzanna. 1996. *Studii od makedonska-bugarska jazična konfrontacija*. [Studies from Macedonian-Bulgarian Contrastive Analysis]. MANU: Skopje.
- Topolińska, Zuzanna. 1997. "Relativisation – a strategy for noun phrase complementation." *Balkanistica* 10. 381-393.

- Topolińska, Zuzanna. 2001. "Zdanie względne – forma czy funkcja?" [Relative clause – form or function?]. *Prace filologiczne*, vol.XLVI. Warszawa. 597-604.
- Topolińska, Zuzanna. 2003. "Means for grammatical accommodation of finite clauses: Slovenian between South and West Slavic". In Janez Oresnik & Donald F. Reindl, ed. *Focus on Slovenian from a typological perspective*. Sprachtypologie und Universalienforschung, Volume 56/3. Berlin: Akademie Verlag. 306-322.