

The Bulgarian Relative Marker *-to**

Catherine Rudin

1. Introduction

Relative pronouns and other relative *wh*-words in Bulgarian are distinguished from interrogative *wh*-words by a suffix, *-to*, as can be seen in the chart in (1).

(1) Bulgarian *wh*-words

| | interrogative (<i>wh</i>) | relative (<i>wh</i> + <i>to</i>) |
|-------------------------|-----------------------------|------------------------------------|
| who (m) | koj | koj <u>to</u> |
| who (f) | koja | koja <u>to</u> |
| who (n) | koe | ko <u>to</u> |
| who (p) | koi | ko <u>to</u> |
| whom | kogo | kogo <u>to</u> |
| what/which/what kind(n) | kakvo | kakvo <u>to</u> |
| which/what kind (m) | kakǎv | kakǎv <u>to</u> |
| which/what kind (f) | kakva | kakva <u>to</u> |
| which/what kind (p) | kakvi | kakvi <u>to</u> |
| whose (m) | čij | čij <u>to</u> |
| whose (f) | čija | čija <u>to</u> |
| whose (n) | čie | čie <u>to</u> |
| whose (p) | čii | čii <u>to</u> |
| where | kǔde | kǔde <u>to</u> |
| when | koga | koga <u>to</u> |
| how | kak | kak <u>to</u> |
| how much/how many | kolko | kolko <u>to</u> |

* A version of this paper was presented at the 2008 Slavic Linguistics Society meeting in Columbus, Ohio. I would like to thank Wayles Browne, not only for his comments at that meeting, but for all the wise suggestions and eye-opening questions he has given me on other work over the decades.

This suffix occurs obligatorily on all and only relative *wh*-words. *Wh*-words without the *-to* suffix are used in questions, both main-clause and embedded, while those with *-to* form headed and headless relative clause constructions. The two are never interchangeable:

- (2) a. **Koj** ima vreme? (***kojto**)
 who has time
 ‘Who has time?’
- b. Ne znaem **koj** ima vreme. (***kojto**)
 NEG know_{1PL} who has time
 ‘We don’t know who has time.’
- c. Čovek, **kojto** ima vreme... (***koj**)
 person who-*to* has time
 ‘A person who has time...’
- d. **Kojto** ima vreme (može da dojde)... (***koj**)
 who-*to* has time can to come
 ‘Whoever/he who has time...’

These facts are straightforward and well known, stated as a simple descriptive generalization in virtually every grammar of Bulgarian. However, even complete reference grammars generally give no attention to *-to* beyond simply noting its existence,¹ and to the best of my knowledge the *-to* suffix has never been discussed from a generative perspective. The goal of this paper is to investigate the status of *-to* in more depth. In section 3 I propose an analysis of relative *-to* as a marker of syntactic status, namely as head of a projection that attracts [–interrogative] *wh*-words to its Spec position. In essence I propose to treat *-to* as a relative complementizer. Section 4 introduces supporting and complicating evidence from two multiple-*wh*- relative constructions, and section 5 discusses some remaining problems. In section 2,

¹ Scatton 1984, for instance, simply states “Relative *pro*-forms are formed from interrogative *pro*-forms with the addition of the postposed particle *-to*” (154). Hauge 1999 is similarly brief: his entire comment on relative pronouns is “Most of them are formed by adding *-to* to the interrogative pronouns” (59) plus a short discussion of the one exception, the stylistically marked (old-fashioned, bookish, and/or dialectal) alternate word for ‘what,’ *što*, which among other idiosyncrasies has no suffixed form.

before introducing the complementizer analysis, I briefly dispense with an appealing but incorrect alternate analysis of *-to* as a marker of definiteness.

2. Rejected Analysis: *-to* Is Not an Article or Definiteness Marker

The *-to* suffix superficially resembles the postpositive definite article; compare (1) and (3):

(3) selo ‘village’ vs. seloto ‘the village’

It is tempting to identify relative *-to* as a definiteness marker, closely related if not identical to the definite article. I have made this assumption myself in earlier work (Rudin 1986), and the idea has a long history in Bulgarian studies, going back at least to Ivanov 1894.² Cyxun 1981 divides Balkan Slavic dialects into those whose relative pronouns are “articled” (*člennie*), that is, suffixed with *-to*, versus “unarticled” (*nečlennie*), with no *-to*. The assumption is reasonable on general linguistic grounds, since question words are presumably indefinite, while relative pronouns can be considered definite; see eg., Bianchi 1999, Bhat 2004, Šimík 2007. Relative pronouns are historically derived from definite determiners or demonstratives in many languages worldwide, indicating a clear crosslinguistic link between relative pronouns and definiteness; see e.g., Andrews 1975. A few examples are given in (4) of languages in which some relative pronouns incorporate at least an etymological definite article or in which the relative pronouns are actually identical to definite articles:

(4) a. *Relative Incorporates Definite Article*

| | |
|------------------|----------------------------------|
| Romance: | lequel, il quale, el cual |
| Modern Greek: | o opios |
| Archaic English: | the which |
| Albanian: | i cili |

² Ivanov is quoted in Cyxun 1981 (108) as saying “I believe that this is the very same *-to* [as the article] albeit invariable in all numbers and genders” (Translation mine).

(4) b. *Relative = Definite Article*German: **der**, etc.Arabic: **al-** ?? (see Belkacemi 1999 for critical discussion)

Bulgarian-internal evidence for a distinction in definiteness between bare and suffixed *wh*-words can be found in the use of bare *wh*-words as indefinite pronouns, as in (5a). This usage is quite restricted even for bare *wh-*, but completely impossible for *wh-to*: see (5b).

- (5) a. Njamam **koj** da mi pomaga.
not-have_{1SG} who to me help
'I don't have anyone (=lit. who) to help me.'
- b. *Njamam **kojto** da mi pomaga.

However, appealing though the idea may be, *-to* is not the definite article. First, *-to* lacks the gender and number agreement of the article, appearing always in what looks like the neuter singular article form, even when accompanying non-neuter *wh*-words. The *wh*-words which are marked for gender or number are repeated in (6).

- (6) **kojto** / **kojata** / **koeto** / **koito** 'who' (relative)
M / F / N / PL
- kakävto** / **kakvato** / **kakvoto** / **kakvito** 'which' (relative)
M / F / N / PL

The invariant ending of these *wh*-words is quite different from the true definite article, which has four distinct forms:

- (7) masculine **-ät** or **a**:³ xljab 'bread' / xljab**ät**, xljaba 'the bread'
feminine **-ta**: kola 'car' / kolab**a** 'the car'
neuter **-to**: kafe 'coffee' / kafet**o** 'the coffee'
plural **-te**: knigi 'books' / knigite 'the books'

³ The two masculine forms are nominative vs. non-nominative case in literary Bulgarian; in colloquial Bulgarian this distinction is lost.

If the relative *-to* suffix were the definite article we would expect unattested forms like those in (8) instead of the actual forms in (6).

- (8) *kojät / *kojata / *koite 'who' (relative)
 M / F / PL
 *kakävät / *kakvata / *kakvite 'which' (relative)
 M / F / PL

Secondly, if *wh*-words did in fact combine with a definite article, we could expect them to occur also with other determiners such as possessives, demonstratives, or the indefinite article *edno*. None of these determiners are actually found with *wh*-words, however. Compare the range of determiners found with a noun in (9a) with the impossibility of any of them in (9b).

- (9) a. deteto mi 'my child'
tova dete 'that child'
edno dete 'a child'
- b. *koj(to) mi 'my who'
 *tova küde 'that where'
 *edno kakvo 'a what'

Finally, no *pro*-forms other than relative *wh*-words occur either with invariant *-to* or with an agreeing definite article, so it would be at least unexpected for relative pronouns to occur with an article. The starred forms in (10) show that personal pronouns, demonstrative pronouns, and adverbial *pro*-forms cannot take an article or the invariant *-to* suffix; this suffix is found only on relative *wh*-words.

- (10) toj 'he' *tojto / *tojät
tja 'she' *tjato / *tjata
onezi 'those' *onezito / *onezite
tam 'there' *tamto
togava 'then' *togavato

It is, in short, abundantly clear that *-to* is not the definite article, regardless of their etymological connection.⁴ One might be tempted to argue that it is still some type of morphological definiteness marking; after all, definiteness can certainly be shown in other ways than through a definite article. However, this also does not hold water: *-to* does not always or reliably indicate definiteness. Even indefinite relative clauses take *-to*; in (11b), for instance, we have *kojto* in a situation where it is not even specific, much less definite; the speaker is referring to any chair, not a certain specific chair.⁵ The *wh*-word must have the *-to* suffix in this example, just as in the definite relative in (11a).

- (11) a. stolăt na **kojto** sedna
 chair_{DEF} on which sat
 ‘the chair on which s/he sat’
- b. stol na **kojto** da sedna
 chair on which to sit
 ‘a chair on which to sit’

3. Proposal: *-to* Heads a Relative Clause Projection

I take it as established that *-to* is not a marker of definiteness. In fact, it appears to be, as the traditional grammars imply, simply a relative clause marker. But what exactly does this mean? I suggest that, rather than being just a bit of morphology attached to *wh*-words, the *-to* suffix has independent syntactic status as the head of a non-interrogative *wh*-clause; that is, it is essentially a relative complementizer.

One hint that this analysis might be right comes from the very closely related Macedonian, whose relative marker, optionally added to most interrogative *wh*-words, actually has the form of a complementizer. The main *wh*-words, interrogative and relative, are shown in (12).

⁴ Relative *-to* and the definite article do derive from the same source, a Common Slavic demonstrative pronoun. See Mladenova 2007 (241–43) for discussion and references on this historical development.

⁵ A reviewer points out that the same is true for French *lequel*, Italian *il quale*, etc. I would argue that this indicates these forms also are not necessarily definite, their etymological connection to the definite article notwithstanding.

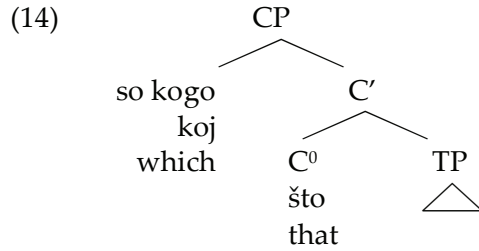
(12) Macedonian *wh*-words:⁶

| | interrogative | relative (<i>što</i> is optional) |
|--------------------|---------------|------------------------------------|
| who, which | koj | koj <u>što</u> |
| whom | kogo | kogo <u>što</u> |
| what kind | kakov | kakov <u>što</u> |
| whose | čij | čij <u>što</u> |
| where | kade/kaj | kade <u>što</u> , kaj <u>što</u> |
| how | kako | kako <u>što</u> |
| how much, how many | kolko | kolku <u>što</u> |
| how big | kolkav | kolkav <u>što</u> |

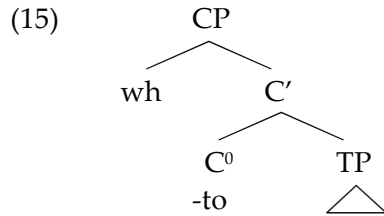
Relative *wh*-words are formed “by adding *što*, written without any space after *koj* and *čij* (hence, *kojšto*), and written as a separate word after all others” (Friedman 2001). This *što* element appears to be the same word as the relative complementizer word *što*, so that a Macedonian relative clause like (13a–b) has the same structure as a Middle English one like (13c). This structure is presumably that shown in (14), with the *wh*-phrases in SpecCP, immediately followed by *što* or *that* in C.

- (13) a. čovekot [so **kogo što** se šetaše včera]
 person_{DEF} with whom that REFL walked yesterday
 ‘the person with whom he walked yesterday’
 (slightly modified from Friedman 2001)
- b. čovekot [**kojšto** mislam deka te videl tebe]
 person_{DEF} whom-that think_{1SG} that you saw you
 ‘the person that I think saw you’
 (slightly modified from Friedman 2001)
- c. I am she [**which that** saved hath youre lyf].
 (Middle English (14th century) from Allen 1980)

⁶ *Wh*-words shown in the table are those which can occur with *što*. *Koga* ‘when’ and *što* ‘what’ do not take the relative ending *što*. Gender and number marked forms do take *što*: *kojašto* ‘who_F’ etc. The conditions under which suffixed *što* occurs are not clear to me; grammars of Macedonian, including Friedman 2001, state simply that it is optional.



Let us suppose that Bulgarian also has a structure like this; that is, let us hypothesize that *-to* is merged in the head of the projection which attracts fronted *wh*-words. Given that Bulgarian is a true *wh*-movement language, this projection is CP. We could see *-to* as spelling out the features which distinguish relative from interrogative *wh*-clauses, features which are located in C. The structure of a Bulgarian relative clause will then be (15), identical to (14) except that the material in C ends up as a suffix and not an independent word.



The claim is thus essentially that *-to* is a complementizer. Not surprisingly, it is in complementary distribution with other complementizers, especially the relative complementizer *deto*, as seen in (16). One possibility would be to see *-to* in fact as an allomorph of *deto*; the relative complementizer appears as *deto* when alone and as *-to* when accompanying a *wh*-word.⁷ Under this analysis, the resumptive pronoun which regularly appears in *deto* relatives (*gi* in (16a)) could be seen as corresponding to the *wh*-portion of the *wh+to* word in (16b) and other *wh*-relatives.

⁷ This is not the source of *-to* historically; in fact, *deto* itself derives from the *wh*-word (with *-to* suffix) meaning 'where': *kūdeto* > *gdeto* > *deto*.

- (16) a. *filmite, deto gi gledahme*
movies_{DEF} that them watched_{1PL}
 ‘the movies that we watched’
- b. *filmite, koito gledahme*
movies_{DEF} which watched_{1PL}
 ‘the movies which we watched’
- c. **filmite, koito deto (gi) gledahme*
movies_{DEF} which that them watched_{1PL}
- d. **filmite, koi deto (gi) gledahme*
movies_{DEF} which that them watched_{1PL}

Other complementizers, declarative *če* and interrogative *dali*, also do not cooccur with *-to*. Not only would they be in competition for a single C slot, they also conflict in clause type.

4. Single and Repeating *-to* in Multiple Relatives

The analysis of *-to* as a clausal head is supported by evidence from a perhaps unexpected quarter: an otherwise odd fact about multiple *wh*-constructions. In spite of the clear pattern we saw in (2), that bare *wh*-words in Bulgarian are normally only interrogative, there is one situation in which bare *wh*-words, without *-to*, can unexpectedly function as relative pronouns. This one anomalous case is multiple *wh*-free relatives and correlatives. In the examples of multiple correlatives in (17) and (18); notice that both of the *wh*-words can have the *-to* suffix, or the first *wh*-word can be unsuffixed.

- (17) a. [**Koj** **kakvoto** iska], da vzeme.
 who what-*to* wants to take
 ‘Let everyone take what they want’ (pair-list)
- b. [**Kojto** **kakvoto** iska] da vzeme.
 who-*to* what-*to* wants to take
 ‘Whoever wants whatever, let them take it’ (single-pair)

- (18) a. [**Kogo kakvoto** go boli], za nego prikazva.⁸
 whom what-*to* him hurts about it talks
 ‘Everyone talks about whatever is hurting them.’ (proverb)
 (pair-list)
- b. [**Kogoto kakvoto** go boli], za nego prikazva.
 whom-*to* what-*to* him hurts about it talks
 ‘The person who has something hurting, talks about it.’
 (single-pair)

These two options are not quite synonymous. The difference between the (a) examples with initial bare *wh-* and (b) with *-to* on both *wh-*words is one of pair-list vs. single-pair interpretation.⁹ So, (17a) refers to a set of people, each of whom wants something, possibly a different thing in each case, while (17b) refers to a single unknown person who wants a certain thing. Similarly (18a) implies multiple people, one with a headache, another with a broken heart, and so on, while (18b) refers to a single individual. These judgments are subtle, but fairly robust.

How does this fact help us come to grips with the syntax of the *-to* suffix? An obligatory pair list reading in multiple questions is well known to correlate with movement of *wh-* to SpecCP. Thus, multiple questions in English (with *wh-*movement to SpecCP) normally elicit a pair-list answer. Single-pair answers are normal only in languages with no overt *wh-*movement, for example, Japanese. Among Slavic languages, single-pair answers are possible in, e.g., Serbo-Croatian, which arguably fronts *wh-*words to a focus position instead of SpecCP in most contexts (see Grebenyova 2006, Bošković 2003, and sources cited there, especially Hagstrom 1998).¹⁰ This suggests that the (a) and

⁸ Some speakers prefer *koj(to)* instead of *kogo(to)* in (18a, b), perhaps for reasons connected with the ongoing replacement of *kogo* with *koj* in colloquial Bulgarian; the remnants of the case distinction are collapsing much as for English *who/whom*. The clitic *go* is obligatory with the experiencer verb *boli* and clitic-doubles the accusative *kogo*.

⁹ A reviewer suggests the distinction could also be seen in terms of indefinite vs. specific interpretation, an idea I leave for further research.

¹⁰ Hagstrom proposes that multiple questions are interpreted via a Q morpheme which is an existential quantifier moved to C; Bošković adopts this

(b) examples in (17–18) differ in the position of the *wh*-words; specifically that the (a) versions have the *wh*-words in SpecCP, while the (b) versions do not. In other words, single and repeating *-to* are in different positions.

This line of thought is immediately reminiscent of the well-known typology of multiple *wh*-movement first sketched out by Rudin 1988 and refined by large numbers of researchers over the last two decades. Some multiple *wh*-fronting languages place their *wh*-words all in SpecCP, while others place only one or even none in SpecCP, fronting the rest to some other left-periphery position such as SpecFocusP or Spec of some other functional projection. There is some evidence that different constructions within a language can have different *wh*-positions. Bulgarian normally (which is to say, in multiple questions) fronts all *wh*-words to SpecCP. Perhaps the construction in (17b) and (18b) shows that Bulgarian is not monolithic in this regard: in this one construction Bulgarian places its *wh*-words in a non-SpecCP position, a configuration more typical of languages like Polish or Serbian/Croatian.¹¹

There are two multiple relative constructions in Bulgarian, multiple free relatives and multiple correlatives. These constructions differ in their external syntax, that is, their relation to the matrix clause. Correlatives are always sentence-initial, external to the main clause and linked to it by a coreferential pronominal or demonstrative, the “correlate.” Free relatives are inside the matrix clause and function as DP or AdvP within it. These characteristics are sketched in (19). (See Rudin 2007a for more detailed discussion.)

proposal and suggests that movement of *wh* over C invalidates the single-pair reading.

¹¹ Just as an aside, the ability to place some or all *wh*-words in a position other than SpecCP may account for the ability for at least some Bulgarian speakers to place adverbs and parentheticals between the *wh*-words of a multiple question, albeit not as freely as in some other languages.

(19) Two types of (multiple) headless relative constructions

| <i>Correlative</i> | <i>Free Relative</i> |
|--|---|
| [[CP wh wh ...] _i [CP ... CORRELATE _i ...]] | [CP ... [DP/AdvP [CP wh wh ...]] ...] |
| <ul style="list-style-type: none"> • adjoined to CP: left peripheral position • not dominated by any XP • coreferential pro-form in main clause | <ul style="list-style-type: none"> • occupies argument or adverbial position • dominated by DP or AdvP • no coreferential pro-form |

Correlatives and free relatives also differ in their ability to have *-to* on one or both *wh*-words. Both patterns (**wh-to wh-to** and **wh wh-to**) are found in multiple correlatives like those in (17–18), but only the second pattern is possible in free relatives. An example of a multiple *wh*- free relative is given in (20), showing that this construction cannot have repeated *-to*.¹² As we might expect, this construction has only the pair-list reading; i.e., the reading which normally goes with a single *-to* at the end of the *wh*-word string.

- (20) a. Vzemajte **koj** **kakvoto** može. (pair-list)
 take who what can
 ‘Let everyone take whatever they can’
 b. *Vzemajte **kojto** **kakvoto** može.

The chart in (21) summarizes possibilities for both kinds of multiple relatives.

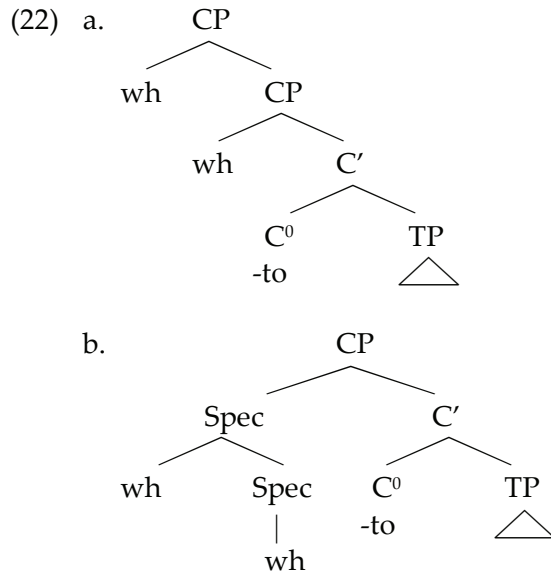
¹² One Bulgarian speaker reports that (20b) is good for her, but most speakers I have consulted do not accept it, and in internet searches I have found the repeating *-to* pattern only in correlatives.

(21) **Possible *-to* configurations, readings, and *wh*-positions in Bulgarian multiple relatives**

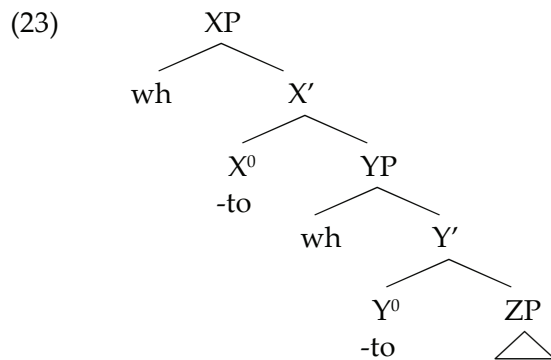
| <i>Multiple Correlative</i> | <i>Multiple Free Relative</i> |
|--|-------------------------------------|
| [CP <i>wh wh-to</i> ...] | [CP <i>wh wh-to</i> ...] |
| pair list <i>wh</i> 's in SpecCP | pair list <i>wh</i> 's in SpecCP |
| [CP <i>wh-to wh-to</i> ...] | _____ |
| single pair <i>wh</i> 's lower than C | |

I have argued elsewhere (Rudin 2007a; Rudin 2008) that the multiple free relative construction (unlike multiple correlative) is possible only in languages which place at least one *wh*-word in SpecCP, while multiple correlatives also occur in languages like Polish which have no true *wh*-movement to SpecCP. This dovetails nicely with the suggestion that single *-to* is possible only with *wh*- in SpecCP.

Returning to the idea that *-to* is a spelling-out of features associated with relative *wh*-, consider how this meshes with the pattern in (21). Suppose that the features spelled out by *-to* are located in whatever functional head(s) attract *wh*-words in a given language and construction, whether C or another projection such as FocusP. When this projection has multiple specifiers, as is the case for CP in languages of the Bulgarian type (the so-called +MFS languages of Rudin 1988), a single instance of *-to* will occur at the end of the string of *wh*-phases occupying those specifiers, as in (22a) or perhaps (22b). Both configurations have been proposed as the structure of multiple fronted *wh*-constructions in Bulgarian (by Richards 2001 and Rudin 1988, respectively). The choice between the two structures in (22), though obviously important for other purposes, is immaterial for the analysis I am proposing here; in either case, both *wh*-words are in a Spec-head relation with a single C⁰, resulting in a surface appearance of *-to* suffixed only to the last *wh*-word. This, I suggest, is the configuration of multiple free relatives and also of multiple correlatives with pair-list reading.



When two instances of the *-to* suffix appear, that is, when each *wh*-word has its own *-to*, the *wh*-words are in Spec of separate projections, and the head of each projection is spelled out as *-to*, as in (23). This is the structure of multiple correlatives with single-pair reading.



Notice that this analysis provides an explanation for the impossibility of *-to* distribution other than on the last of a group of *wh*-words or on each *wh*-. In particular, a multiple relative cannot have *-to* only on the first *wh*- or no *-to* at all; compare (24) to (16).

- (24) a. ***[Kojto kakvo iska]**, da go vzeme.
 who what wants to it take
 b. ***[Koj kakvo iska]**, da go vzeme
 who what want to it take

Assuming that C (and the “X” and “Y” heads in (22), on which more shortly) cannot be null, i.e., that the relative *wh*-attracting heads are obligatorily spelled out as *-to*, and given that specifiers are to the left of the head, there will always at least be a *-to* at the end of the *wh*-string.

5. Remaining Problems

Although the analysis of *-to* as head of a *wh*-attracting projection accounts nicely for several aspects of its syntax, including the single vs. repeating *-to* of multiple relative constructions, it also raises some issues which remain unsolved at this point. The first such problem is the identity of the projections labelled XP and YP in (23). One of these could be a focus projection; *wh*-fronting is often said to be attraction to Focus. Perhaps the other is TP or perhaps another discourse-related category. I leave this as an open question for now; what matters for this analysis is that they must be two separate projections with separate heads. One other point which is clear is that neither XP nor YP can be CP for Bulgarian sentences like (17b) and (18b), since movement of even one *wh*- to CP should force the pair-list reading.¹³

A second unresolved issue raised by this analysis of *-to* is how to deal with *wh*-phrases like those in (25). In such multi-word *wh*-phrases *-to* is suffixed to the *wh*-word, not to the whole phrase, as we might expect if the *wh*-phrase occupies the Spec position of the head that is spelled out as *-to*.

- (25) a. **kojato** kniga (***koja knjigato**)
 ‘which book’

¹³ However, it is possible that languages other than Bulgarian might have a multiple *wh* correlative structure with one *wh* in SpecCP (and pair-list reading). I have not investigated interpretation of multiple correlatives in other multiple fronting languages.

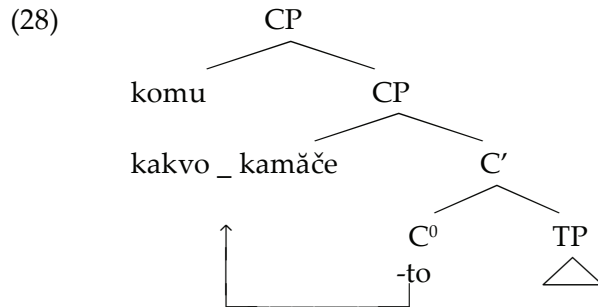
- (25) b. **kakvoto** kamăče (***kakvo kamučeto**)
 ‘what kind of little stone’

This position of *-to* is reminiscent of the position of other DP clitics in Bulgarian, namely, the definite article and possessive clitics, which also suffix to a modifier rather than the entire phrase. This is essentially a type of second-position cliticization within the nominal phrase. (The precise rule for placement of these clitics is more complex, but need not concern us here.)

- (26) a. novata kniga (*nova knjigata)
 ‘the new book’
 b. mladoto ti dete (*mlado deteto ti)

Apparently a similar second-position placement applies in relative *wh*-phrases, though the exact mechanism for raising *-to* into the *wh*-phrase remains to be explored. Phrase-internal *-to* occurs also in multiple relatives, in sentences like those in (27). Under the present analysis these must have a structure something like (28), with *-to* somehow cliticized into the 2nd position of the preceding *wh*-phrase.

- (27) a. Săsedite bjaha si otmŭkvali **komu** **kakvoto**
 neighbors_{DEF} had REFL carried-off whom_{DAT} which
 kamŭče potrjabvalo.
 stone_{DIM} was.necessary
 ‘The neighbors had carried off whichever little stone each
 one needed’ (Daskalov)
 b. Vzemi otkăde kojata kniga iskaš.
 take from-where which book want_{2SG}
 ‘Take whatever book you want from wherever.’



A cliticization analysis of some kind is independently supported by the fact that *-to* behaves like a clitic in terms of morphology in any case. As noted by Pashov 1999, it attaches to the end of the fully inflected word, outside of all gender, number, and case inflections:

- | | | | |
|------|-----------|-------------------------|---------------|
| (29) | koj+a+to | which _F | *koj+to+a |
| | kakv+i+to | what.kind _{PL} | *kav(ă)v+to+i |
| | ko+go+to | whom | *ko(j)+to+g |

This morpheme order is highly unexpected if *-to* were a derivational suffix; in that case it should attach to the root with inflectional affixes further out, as in the starred forms in (29). But attaching to the fully inflected word is entirely normal for a clitic. It is thus not surprising that it has the same type of syntactic mobility as other phrase-internal clitics, i.e., roughly 2nd place within its phrase, even if the apparent movement required is somewhat problematic.

It is worth noting that the same problem applies in Macedonian, where *što*, if present, follows the *wh*-word in complex *wh*-phrases just as *-to* does in Bulgarian:

- (30) a. ...**kakov** **što** **čovjek** ...
 ...what.kind that man
 'which(ever) kind of man'
- b. ...katolici **vo** **čij** **što** **crkvi** služat...
 ...catholics in whose that churches serve_{3PL}
 'catholics in whose churches they serve'

[archives.vmacedonia.com/15285.htm]

If Macedonian *što* is actually the complementizer *što*, examples of this type pose the same conundrum as the Bulgarian phrases in (25): apparent movement of an element in C into a *wh*-phrase. One possibility is that in both languages the *wh*-phrase moves first to a position below C, perhaps SpecFocusP, and that the *wh*-word alone then moves further to SpecCP; this idea has problems of its own, however, and I leave the issue open for further research.

6. Conclusion

In this paper I have suggested an analysis of relative *-to* as a marker of syntactic status. Specifically I suggest that *-to* spells out features which identify a relative clause, perhaps something like [–interrogative, +wh]. These features are located in the head of projections which can attract the relative *wh*-word or phrase. In Bulgarian this projection is normally C: *wh*-words, including all the *wh*-words in a multiple-*wh*-construction, normally move to SpecCP. Thus *-to* is essentially a relative complementizer. Fronted *wh*-words in one construction, the multiple correlative, do not necessarily move to SpecCP, and those cases where they front to a non-CP projection are marked by the occurrence of multiple *-to* markers, one per projection (i.e., each head that attracts a relative *wh*-word is spelled out separately as *-to*).

Questions remain, obviously. One problem is the identity of the heads/projections involved in the multiple *-to* construction. Another is the mechanism for placing *-to* correctly in multi-word *wh*-phrases. But in spite of these issues I find the analysis appealing enough to be worth pursuing.

Some of the results of the paper hold regardless of whether the relative complementizer analysis turns out to be correct. In the first part of the paper I demonstrated that *-to* is clearly not an article, or even a definiteness marker. It is also clear that it is a clitic, and not a derivational suffix. Finally, whether we actually call it a complementizer or not, it clearly has connections to the complementizer *što* in Macedonian.

To end with a broader view, I close by noting that overlap of determiners and complementizers is known from languages as far afield as Siouan and other Native American language families; clause-typing morphology tends to merge with, develop from or into, and generally be hard to distinguish from nominalizing morphology and especially

definite determiners. So it is not surprising typologically that a particle which began life as a demonstrative and may at one stage have been an article seems to have turned into a complementizer or at least some kind of clause typing marker.

Wayne State College
carudin1@wsc.edu

References

- Allen, Cynthia. (1980) *Topics in diachronic English syntax*. New York: Garland Publishing.
- Andrews, Avery Delano. (1975) *Studies in the syntax of relative and comparative clauses*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Belkacemi, Cemi. (1999) "A view on relativisation in Arabic". *Philologie im Netz* 7: 1–10.
- Bhat, D. N. Shankara. (2004) *Pronouns: A crosslinguistic study*. Oxford: Oxford University Press.
- Bianchi, Valentina. (1999) *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.
- Bošković, Željko. (2003) "On the interpretation of multiple questions". *Linguistic variation yearbook* 1: 1–15.
- Cyxun, Gennadij A. (1981) *Tipologiĉeskie problemy balkanoslavjanskogo jazykovogo areala*. Minsk: Nauka i tehnika.
- Friedman, Victor. (2001) "Macedonian". [Available at: <http://www.seelrc.org:8080/grammar/mainframe.jsp?nLanguageID=3>]
- Grebenyova, Lydia. (2006) "Sluicing puzzles in Russian". James Lavine, Steven Franks, Mila Tasseva-Kurktchieva, and Hana Filip, eds. *Formal approaches to Slavic linguistics: The Princeton meeting*. Ann Arbor, MI: Michigan Slavic publications, 157–71.
- Hagstrom, Paul. (1998) *Decomposing questions*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Ivanov, M. (1894) "Prinos kǎm izučavanie bŭlgarskite dialekti (zabelieŭski po govorite v oblastta na srednogorskoto nareĉie)". *Periodičesko spisanie na bŭlgarskoto kniŭovno družestvo v sredeĉ* 45: 399–418.

- Mladenova, Olga. (2007) *Definiteness in Bulgarian: Modeling the processes of language change*. Berlin: Mouton de Gruyter.
- Pašov, Petăr. (1999) *Bălgarska gramatika*. Plovdiv: Hermes.
- Rå Hauge, Kjetil. (1999) *A short grammar of contemporary Bulgarian*. Columbus, OH: Slavica.
- Richards, Norvin. (2001) *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.
- Rudin, Catherine. (1986) *Aspects of Bulgarian syntax: Complementizers and wh- constructions*. Columbus, OH: Slavica.
- . (1988) "On multiple questions and multiple *wh*- fronting". *Natural language and linguistic theory* 6: 445–501.
- . (2007a) "Multiple *wh*- relatives in Slavic". Richard Compton, Magdalena Goledzinowska, and Ulyana Savchenko, eds. *Formal approaches to Slavic linguistics: The Toronto meeting*. Ann Arbor, MI: Michigan Slavic publications, 282–307.
- . (2007b) "Multiple vs. single *-to* in Bulgarian multiple *wh*-relatives". Presented at SEESA meeting, Columbus, OH.
- . (2008) "Multiple *wh*- non-interrogatives". Presented at University of Wisconsin and University of Minnesota linguistics colloquia, April 4.
- Scatton, Ernest. (1984) *A reference grammar of modern Bulgarian*. Columbus, OH: Slavica.
- Šimík, Radek. (2007) "The sources and consequences of *wh*-morphology in (English) relatives and questions". Presented at the Conference for Students of Linguistics in Europe (ConSOLE), Brussels, January 13.