Pair-List vs. Single Pair Readings in Multiple Wh Free Relatives and Correlatives

Catherine Rudin
Wayne State College

1. Introduction

This paper has two goals. The first and more general one is to describe two different multiple wh relative constructions, the Multiple Wh Free Relative and the Multiple Wh Correlative. Multiple Free Relatives in particular are a rare and poorly understood construction, found to the best of my knowledge only in several Balkan languages. The second, more specific goal is to investigate the distribution of pair-list and single-pair readings in these two constructions and the likely relation of these readings to their structure; this portion of the paper draws heavily on insights provided by Bulgarian wh-word morphology. I conclude that pair-list and single-pair readings are related to the landing sites for multiple fronted wh words in given constructions and given languages.

2. Free Relatives and Correlatives

Before looking at any multiple wh constructions, it will be useful to clearly delineate two types of “headless” relative clauses. The two constructions under consideration in this paper are Free Relatives (FR) and Correlatives (CR). Both are non-interrogative wh clauses, with a fronted wh word (or, as we shall see, wh words). However, in spite of their similarity, they are in fact distinct constructions, differing both in their category status (nominal or not) and their relation to the matrix clause. For clear treatments of the differences between the two constructions, see Izvorski 1996, Citko 2006.

A Free Relative, as shown very roughly in (1a), is a clause which has the internal structure of a relative clause, with a fronted wh word, but instead of modifying a noun it constitutes a nominal phrase all by itself and can function as an argument of a higher clause. In the English and Bulgarian examples in (1b-c) the bracketed clauses are direct objects.

(1) Free Relative (FR)
   a. \[NP\{CP [wh ... ]]\]
   b. I bought \[NP\{CP what was available \} \]
   c. Kupih \[NP\{CP kakvoto imaše \} \] (Bulgarian FR)
      bought(1s) what there-was

There is a long-standing debate about the precise structure of FR, which is glossed over in the sketchy bracketing in (1); in particular, the question of whether the dominating NP has a null head, no head, or raises the wh-word to its head position is one that has been debated since the 1970s.\(^1\) I will have some comments relevant to this issue later in the paper.

\(^1\) Under the Head Hypothesis (e.g. Bresnan & Grimshaw 1978), the wh word is head of the higher nominal phrase; under the Comp Hypothesis (e.g. Groos and van Riemsdijk 1981) wh is in the same position as wh in headed
The second construction of interest, Correlatives, are like FR in being a non-interrogative wh clause with no overt head; they differ from FR in that they are not arguments and presumably not nominal either; the bracketing (2) assumes they are simply CP, though it is possible there may be a dominating phrasal category of some sort. As sketched in (2a), they occur at the far left edge of the main clause, not inside it, and they always have a coreferring demonstrative or pronominal within the main clause, underlined in examples (2b,c):

(2) Correlative (CR)
   a. [CP wh ... ]] [ ... dem/pro ... ] ..
   b. [CP Kojto se uči], [toj šte spoluči]
      who refl studies he will succeed
   c. [CP Whoever studies], [he, will succeed]

Correlatives are awkward in English but grammatical in many other languages, including (perhaps all) Slavic languages. They have received most attention in South Asian languages, but I give a Bulgarian example since this is the language I will be most concerned with in the paper. The examples in (1) and (2) have a single wh word. In the rest of the paper we will be dealing with the same two types of clauses, but with multiple wh words.

3. Both FR and CR can have multiple wh: Bulgarian/Macedonian/Romanian vs. Polish

In at least some Slavic and Balkan languages, multiple fronted wh also occurs in the two “headless” relative constructions, FR and CR. These multiple wh relative constructions have, however, received very little attention, unlike multiple wh questions, which have been very widely discussed in the linguistic literature. Most of the work on multiple wh relatives is my own. Rudin 1986 (chapter 6.3) gives a number of examples and suggests that they may have implications for the analysis of both multiple wh questions and free relatives crosslinguistically. In a series of more recent papers I treat various aspects of both constructions (Rudin 2006, 2007a, 2007b, 2007c). Multiple-wh relative CR have been mentioned in passing in a few works (Williams 1986, Izvorski 1997, 2000, for instance) but multiple FR have gone almost entirely unnoticed, or even implied to be impossible.

The first work to specifically compare FR and CR in terms of multiple wh fronting is Citko 2006. On the basis of Polish data, Citko states that multiple wh is possible ONLY in CR, not in FR; in fact, she presents this as a major factor distinguishing CR and FR. Examples (3a-b) illustrate the contrast. In (3a) a clause with two wh words, kto co, is fine as a correlative (notice it is at the left edge of the sentence, and the main clause contains the typical correlative pronominals). An attempt to use the same clause as a FR in (3b) (in an argument position; not left peripheral; no coreferring pronominal) results in strong ungrammaticality.

relatives, and the head position of the dominating phrase is either null or simply nonexistent. Various versions of both of these hypotheses are still argued for in current work on free relatives.

2 For ease of interpretation of examples the following typographical conventions are used throughout the paper: wh words are **boldfaced**; correlative pronominals are **underlined**.

3 In fact, correlatives are usually said to be ungrammatical in English, but sentences like (2c) do not seem terribly bad to me.
(3)a. [Kto co chce], ten to dostanie. (Polish CR)
   who what wants that this gets
   ‘Whatever anyone wants, they get it.’ (Citko 2006)

   b. *Dostanie [kto co chce]. (Polish *FR)
      gets who what wants
      ‘Everyone gets what they want.’ (Citko 2006)

However, this restriction of multiple wh to CR, while apparently valid for Polish and some other Slavic languages, does not hold for several Balkan languages. Multiple wh phrases do occur in FR as well as CR in Balkan Slavic (Bulgarian and Macedonian) and in Balkan Romance (Romanian). Examples from all three languages are given in (4) through (6). In each pair the (a) example is CR (left peripheral and with correlative pronominal) and the (b) example is FR (non-left-peripheral and with no coreferring demonstrative or pronoun in the main clause). The FR examples are fully grammatical, in stark contrast to Polish (3b).

(4)a. [Na kojto kakvoto e pisano], tova šte stane. (Bulgarian CR)
   to who what is written that will happen
   ‘Whatever is fated for each person, that will happen.’ (web)

   b. Vzemajte [koj kakvoto može]. (Bulgarian FR)
      take who what can
      ‘Let everyone take whatever they can.’ (Mantov)

(5) a. [Na koj što mu e pišano], toa ke se sluči. (Macedonian CR)
   to who what him is written that will refl happen
   ‘Whatever is fated for each person, that will happen.’ (E. Petroska p.c.)

   b. Neka kažuva [koj što saka]. (Macedonian FR)
      let say who what wants
      ‘Let everyone say whatever they want.’ (E. Petroska p.c.)

(6) a. [Cine ce vrea], aia să facă! (Romanian CR)
   who what wants subj. do
   ‘Let everyone do whatever they want.’ (M. Irimia, p.c.)

   b. Trâncânește [cine ce vrea]. (Romanian FR)
      blabs who what wants
      ‘Everyone's blabbing whatever they want.’ (V. Hill p.c.)

Exactly what makes multiple wh FR possible in Bulgarian, Macedonian, and Romanian (but not Polish) is obviously an important question. Indeed, it is a sort of umbrella question covering a host of thorny, unresolved issues in both the syntax and the semantics of multiple wh FR, all of

---

4Romanian speakers vary in their willingness to accept multiple wh CR like this. Though some speakers find them fine, others say they are stilted, “sounds like a proverb,” etc. Interestingly, multiple wh FR seems to be much more clearly and universally acceptable.
which one would hope can be attributed to some deep-down difference between the two types of languages. And in fact I think this can be done, though problems remain.

In this paper I focus on just one corner of the problem, namely, the distribution of pair-list and single-pair readings in multiple wh relative constructions. The possibility for one or the other reading turns out to differ across languages and across the two types of headless relative constructions, in ways which may shed light both on the difference between Polish and the Balkan language and on the structure of multiple wh FR and CR.

4. Pair-List vs. Single-Pair Readings

Once again, we need a little background before continuing. It is well known that multiple questions in various languages can have two types of answers, known as “single-pair” and “pair list.” A question like (7) in English normally has a pair-list reading; that is, it elicits as its answer a list of people and what each one of them bought. A single-pair answer (a single person and a single thing bought) is much more restricted in English. The answer “Stan bought shoes” seems odd for (7) outside of echo or quiz-show contexts. However, in other languages, for example Japanese and Serbo-Croatian, the single-pair answer is normal. (See Grebenyova 2006 and sources cited there)

(7) **Who** bought **what**?
   pair-list: Stan bought shoes, Jen bought a plant, Lorraine bought bagels...  
   ???single-pair: Stan bought shoes.

Multiple wh relative constructions similarly can be judged as having one or the other type of interpretation, and interestingly, it seems that FR and CR differ in their ability to have pair-list and single-pair readings. Specifically, in all cases I have been able to check, multiple wh FR in the Balkan languages have only the pair list reading, while CR can at least sometimes have both readings.

Thus, Bulgarian (8) can only mean that one person will send a certain amount, another person a different amount, and so on. Similarly, Romanian (9) only has the reading that various people should eat various things. Neither of these sentences can be interpreted as suggesting that one certain person (whoever that person may be) should send or eat one certain (unknown) thing.

(8) Praštajte **koj kolkoto** može.  
   send who how-much can  
   ‘Everybody send as much as you can’ (web)  
   (i.e. X send $5, Y send $100, Z send $20...)

On the other hand, in Bulgarian (10) both types of readings are possible with slightly different CR clauses. The (a) version means that someone got something, and whoever he was, he deserved it. The (b) version means that various people got various things and they all deserved
them. Similar facts appear to obtain in Romainan. Speakers I consulted accept (11) with pair-list interpretation, and some also consider a single-pair possible.

(10) a. **Kojto kakvoto** e polučil, si go e zaslužil. (Bulgarian CR)
    who what has received, refl it has deserved
    ‘Whoever got something deserved it.’ (web)

    b. **Koj kakvoto** e polučil, si go e zaslužil. (Bulgarian CR)
    who what has received, refl it has deserved
    ‘Everyone deserved whatever they got.

(11) **Cine ce** vrea, aiia să facă! (Romanian CR)
    who what wants that subj. do
    ‘Let everyone do whatever they want.’

    ?‘Whoever wants to do something, let him/her do it.’

What could explain the difference in readings between the two types of multiple wh relatives? Why should FR have only Pair-List, while CR can have both readings? A first step toward an explanation may involve the morphology of wh words in Bulgarian.

5. Single vs. Multiple -to in Bulgarian

In Bulgarian, pair-list vs. single-pair readings are signaled by a morphological difference in the wh words themselves. We have already seen a hint of this in (10), where the suffix -to occurs on both wh words in the single-pair (a) version but only on the second wh word in the pair-list (b) one. This judgement is subtle but consistent among many Bulgarian speakers. A couple more examples of this phenomenon are given in (12) and (13).

(12) a. **Kojto kakvoto** iska, da go vzeme. (Bulgarian CR)
    whoDEF whatDEF wants to it take
    ‘Whoever wants something should take it.’

    b. **Koj kakvoto** iska, da go vzeme. (Bulgarian CR)
    who whatDEF wants to it take
    ‘Let each person take whatever they want.’

(13) a. **Kogoto kakvoto** boli, za nego prikazva. (Bulgarian CR)
    whom what hurts about it talks
    ‘The person who has something hurting, talks about it.

    b. **Kogo kakvoto** go boli, za nego prikazva (Bulgarian CR)
    whom what him hurts about it talks
    ‘Everyone talks about whatever is hurting them.’ (proverb)

The correlation of repeated -to with single-pair interpretation is useful in two ways. First, from a practical point of view, it provides an easy way to recognize single-pair and pair-list readings in written sources (especially the internet) without having to interpret context or badger speakers

---

5 The situation is probably complicated somewhat by dialect and possibly stylistic factors, since not all dialects have the -to suffix on relative pronouns.
for difficult, subtle judgements. A search of attested examples confirms that CR and FR differ dramatically with respect to repeating vs. single -to. Specifically, CR occur with both patterns, while clear FR always have single -to.\(^6\) A typical FR example is (14a)(=4b)). An attempt to form a hypothetical FR with repeated -to suffixes as in (14b) results in ungrammaticality. Such sentences are not found in texts and are judged deviant by the speakers I’ve asked.

(14)  
\[  
\text{a. } \text{Vzemajte } \text{koj} \text{ kakvoto može.}  
\text{take who what can}  
\text{‘Let everyone take whatever they can’ (pair-list)}  
\text{b. * Vzemajte kojto kakvoto može.}  
\]

In other words, it is confirmed that FR are always pair-list, while CR can also be single-pair. The patterns we’ve seen so far can be summarized as in (15). Pair-list readings are found in both FR and CR, both with the single -to suffix, while single-pair readings are found only in CR, with the repeating -to suffix.

(15)

<table>
<thead>
<tr>
<th></th>
<th>pair-list</th>
<th>single-pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR</td>
<td>ok (wh wh-to)</td>
<td>*</td>
</tr>
<tr>
<td>CR</td>
<td>ok (wh wh-to)</td>
<td>ok (wh-to wh-to)</td>
</tr>
</tbody>
</table>

Secondly, in addition to making it easy to spot readings, single vs. multiple -to may provide a key to explaining WHY the different interpretations differ, and may also elucidate the structure of multiple wh relatives of different types and in different languages. To see why this is true, we need a little more background on the morphology of wh words in Bulgarian. Normally, bare wh words are interrogative, while wh words with the -to suffix are relative pronouns. So in the single-wh examples in (16) the (a) and (b) sentences are questions, either main clause or embedded, and can only have bare koj, while the (c) and (d) examples are relative clauses (one ordinary headed relative and one FR) and both can only have kojto.

(16)  
\[  
\text{a. Koj ima vreme? (kojto)}  
\text{who has time}  
\text{‘Who has time?’}  
\text{b. Ne znam koj ima vreme. (kojto)}  
\text{neg we-know who has time}  
\text{‘We don’t know who has time.’}  
\text{c. Čovek, kojto ima vreme... (koj)}  
\text{person who has time}  
\text{‘A person who has time...}  
\text{d. Kojto ima vreme (može da dojde) (koj)}  
\text{who has time can to come}  
\text{‘Whoever/he who has time ...}  
\]

\(^6\) There are some non-interrogative multiple wh clauses that are not clearly FR or CR (eliptical main clauses, adverbial or concessive clauses.) I leave these for future research, but it appears likely that they will eventually fit into the analysis proposed here. In the data I have collected, the eliptical main clause type all have single -to, like FR. while the adverbial/concessive type is found with both single and repeating -to, like CR.
The -to suffix thus has the function of making interrogative wh into a relative pro-form. However, in multiple relatives, the simple distinction between bare wh interrogative and suffixed wh relative breaks down; the bare wh in (14a) (or any of the examples earlier in the paper) is clearly not interrogative. It appears that a single -to suffix is at least in some cases sufficient to mark a whole group of wh words relative. This suggests\(^7\) that the multiple wh words form a single constituent, as in (17a), and furthermore that the multiple whs with -to on each wh might have a different structure, in which each wh is a separate constituent, as in (17b).

\[(17)a. \ [\text{wh wh}]\text{-to.} \]
\[b. \ [\text{wh } ]\text{-to [ wh ]\text{-to} }\]

6. A Hypothesis: Pair-list reading requires +MFS structure.

Let us hypothesize, then, that the pair-list and single-pair multiple wh relatives differ in syntactic constituency. In particular, suppose that the two wh words in the pair-list construction are both in SpecCP, while in the single-pair interpretation construction (with repeated -to in Bulgarian), they are in separate Spec positions, perhaps SpecCP and something like SpecFocP. I have labeled the second Spec “FP” in (18) with deliberate vagueness; it could be F=Focus or F=any functional category.

\[(18)a. \ [\text{SpecCP wh wh }\text{-to} \]
\[b. \ [\text{SpecCP, wh }\text{-to [ SpecFP, wh ]\text{-to}} \]

This is immediately reminiscent of the typology of wh-fronting in multiple questions that has been around since Rudin (1988), with refinements by numerous linguists, especially Bošković (1997 and subsequent works), among many others. The typology is sketched in (19). Note that the crucial structural difference, just as in (18), is whether fronted wh words all cluster in SpecCP or are separated into two (or possibly more) constituents.

\[
\begin{array}{|c|c|}
\hline
\text{“+MFS”} & \text{“-MFS”} \\
\hline
\text{all wh words land in SpecCP} & \text{at most one wh word in SpecCP} \\
\quad \quad [[\text{SpecCP wh wh }[C^\prime \ldots]]] & \quad \quad [[\text{SpecCP wh }[C^\prime[F\text{P, wh}] \ldots] \\
\quad \quad \quad \quad \text{or } [C^\prime[F\text{P, wh} ][F\text{P, wh}] \ldots]] \\
\text{Bulgarian} & \text{Polish} \\
\text{Romanian} & \text{other Slavic languages} \\
\text{Macedonian} & \\
\hline
\end{array}
\]

Stripped to its somewhat oversimplified essentials, this typology divides multiple wh fronting languages (that is, languages in which all wh phrases are moved to the left edge of the clause) into two types, with different landing sites for the fronted wh words. In one group of languages, which I follow the original (1988) paper in labelling “+MFS” languages, all wh phrases undergo true wh-movement to SpecCP. This group includes Bulgarian and some other Balkan languages.

\(^7\) As I noted already in my dissertation (published as Rudin 1986).
In the other, “-MFS” group, some or perhaps all of the wh words are fronted not to SpecCP but to other left-peripheral positions such as Focus. This second group includes Polish and other Slavic languages such as Serbo-Croatian. Bulgarian, Romanian, and Macedonian share a list of traits including obligatory fronting of all wh words, multiple long-distance extraction, superiority effects, lack of wh-island effects, and resistance to splitting of the group of wh words, all of which have been attributed to this structural parameter.

The parallelism of (18) and (19) strongly suggests that the ±MFS parameter is involved in single-pair vs. pair-list interpretation. Since the typological split was originally proposed, it has become clear that languages are not necessarily monolithic in their orientation to this parameter; within a given language it is possible for constructions to differ in their wh landing sites. For instance, it has been shown in Serbo-Croatian, Slovene, and Russian multiple questions that under conditions of sluicing or long distance extraction characteristics of true wh movement to SpecCP emerge, while wh in other constructions has the characteristics of movement to a non-SpecCP position. (Bošković 1997, Golden 1997, Grebenyova 2005, and others.) What I am proposing here is essentially another example of what we might call the “split MFS” phenomenon: Bulgarian, though basically a +MFS language, has at least one construction with a non-MFS-like structure, the CR with pair-list interpretation and repeated -to affix would go in the right-hand column of chart (19). To put it another way, a revised version of (15) could be (20), with pair-list identified with +MFS (all wh words in a single specifier) and single-pair with -MFS (wh words not forming a single constituent). Bulgarian constructions with single -to are +MFS structures, while repeated -to constructions have a -MFS structure.

Let us hypothesize further that the ±MFS parameter also accounts for the distribution of multiple wh FR across languages, that is, it accounts for the existence of multiple wh FR in Bulgarian and two other Balkan languages but not in Polish. Under this view it is no coincidence that Bulgarian, Macedonian and Romanian group together in having multiple wh FR. They are precisely the three languages which differ from Polish and most other Slavic languages in the landing site of fronted multiple wh words. If the ability to have more than one wh word in SpecCP is what somehow licenses multiple wh FR, it makes sense that these same three languages are the ones with this construction.

This hypothesis implies a further prediction: multiple wh FR should only be possible in multiple wh-fronting languages, since languages which have only single wh-movement (like English) or no wh movement (like Chinese) would never have the requisite configuration of wh words in a single specifier position for multiple wh FR.

---

8 The availability of -MFS structures in Bulgarian might also explain the ability for some speakers to insert adverbs within the wh-word string under some circumstances, discussed for instance by Lambova 2003.
It might seem that a second prediction is implied as well, namely that if multiple wh CR
are possible in any non-multiple wh-fronting languages, they should have only the single-pair
interpretation, since, those languages again could not have the necessary configuration for a pair-
list reading. However, since pair-list readings are clearly possible without +MFS structure in
multiple questions in non-multiple-fronting languages (such as English; see (7) above), this
second prediction is clearly too strong. Instead, the correct generalization is likely to be that
single-pair readings in mult wh CR in all types of languages are available only when none of the
wh words are in SpecCP. Bošković (2002 and elsewhere) claims that overt wh movement to
SpecCP forces a pair-list reading in multiple questions. Given this it’s not surprising that single-
pair readings in Bulgarian are available only in the construction in which the wh words are not
all in SpecCP. Presumably the Bulgarian multiple wh CR with single pair readings have neither
wh word in SpecCP.

The idea that +MFS structure is necessary for multiple wh FR is quite strongly supported
by empirical facts, namely the correlation of multiple wh FR with single -to in all attested
elements. However, to this point no explanation has been offered for exactly how or why +MFS
structure licenses multiple wh FR. I propose that the explanation has to do with the fact that FR,
unlike CR, are nominal phrases, and the wh words in them must be construed with the head of
the nominal construction. The wh words cannot themselves be heads of the construction.
Having multiple heads poses insurmountable syntactic and semantic problems, and having one of
the whs as head gives the wrong semantics. Multiple wh FR thus clearly support a “Comp
Hypothesis” view of FR structure, and furthermore a version of the Comp Hypothesis involving
a null head of the higher nominal phrase. Let us suppose that a null head is able to be construed
with multiple wh phrases in a single SpecCP, but not with wh’s in separate Spec positions. This
would mean that, of the three configurations in (21), only the first one -- the +MFS one -- is
grammatical.²⁹

(21) \[\begin{array}{l}
\sqrt{[\text{NP e } ][[\text{SpecCP wh wh }], \ldots ]} \\
* [\text{NP e } ][[\text{SpecCP wh }], [C \text{ wh }], \ldots ] \\
* [\text{NP e } ], [\ldots [\text{FP wh }], [\text{FP wh }], \ldots ]
\end{array}\]

CR, on the other hand, are not noun phrases, have no head, and obviously have no requirement
for the wh words to be construed with any head. Therefore, they are free to have one or both of
the wh words in a position other than SpecCP. Multiple wh CR in different languages, or indeed
different types of multiple wh CR within one language (as we have seen in Bulgarian) can have
their wh words either in SpecCP or not.

These ideas are obviously still somewhat vague and speculative, and theoretical details of
the analysis remain to be worked out. But the overall approach, the general outline of the
analysis, does seem quite well supported by data from an increasing number of languages.

8. Conclusion

I take it as established that multiple wh relative constructions differ in wh landing sites from one
language to another in ways that parallel the structure previously established for multiple wh
questions. (Parallelism of Superiority effects in questions and relatives across languages and across

²⁹ Note that it seems a lexical head, unlike a null one, cannot be construed with multiple wh’s under any
circumstances; as far as I know no language has overtly headed multiple relative clauses.
constructions also supports this conclusion, though I didn’t go into it here). The correlation of multiple FRs with +MFS languages strongly suggests that the ability for multiple wh phrases to land in SpecCP is crucial to the grammaticality of multiple wh FR. This conclusion is reinforced by the fact that multiple wh FR in Bulgarian are found only with the single -to pattern (which I have claimed indicates +MFS structure, with all wh phrases forming a single constituent).

Furthermore, I claim that the differing availability of single-pair and pair-list readings between correlatives and free relatives is attributable to the same structural parameter. Single-pair readings are impossible with wh in SpecCP, so are ruled out in +MFS structures. This makes pair-list the only possible reading for multiple wh FR. Multiple wh CR, on the other hand, are able to occur with both pair-list and single-pair readings, and with the corresponding morphological forms in Bulgarian (single -to and repeating -to suffix patterns). These generalizations are predicted to hold universally, meaning that only multiple wh-fronting languages with +MFS structure will have multiple wh FR, always with pair-list semantics, while a variety of languages may have multiple wh CR.

References


Rudin, Catherine (1986). *Aspects of Bulgarian Syntax: Complementizers and wh Constructions.* Slavica, Columbus, OH.


Rudin, Catherine (2007c). “Multiple vs. Single -to in Bulgarian Multiple wh Relatives” Paper presented at SEESA meeting, Columbus, OH

Williams, Joanna Radwańska (1986). Wh constructions and underlying word order in Polish. Paper presented at AATSEEL annual meeting.

**Author contact information:**

*Catherine Rudin: carudin1@wsc.edu*