Relativization, Intonational Phrases and Rich Left Peripheries

Simona Herdan
University of Connecticut
Relativization, Intonational Phrases and Rich Left Peripheries
Relativization, Intonational Phrases and Rich Left Peripheries

Simona Herdan*

1 Introduction

The goal of this paper is to provide an account of some novel data involving relativizer restrictions in some of the Slavic languages and in Romanian. In doing that I will discuss the consequences of these facts for the syntax of the complementizer area of the clause. I will conclude that the relativizer restrictions provide evidence for locating the relevant relativizers, roughly corresponding to that and which in English, not just in different places in the same phrase, but in different phrases altogether (cf. Bianchi, 1999). I will also show that the need to align intonational phrasing with the syntax, as well as the structure of the nominal domain, i.e., the presence or absence of D, is responsible for the observed restrictions.

The paper will be organized as follows. In section 2, I will present the paradigm to be accounted for and a first generalization, followed by an outline of the proposal. In section 3, I will present evidence for the role of intonational phrasing in these relativizer restrictions and discuss the implications for the syntax of relativization. Finally, section 4 will show how we can derive the relativizer restrictions (or the absence thereof) and discuss the role of nominal projections. Section 5 will contain the conclusions.

2 The Data and a Proposal

2.1 The Basic Paradigm

Languages that possess more than one relativizer generally allow them to be used interchangeably, as illustrated in (1) for English.

(1) John got the present that/which/∅ he wanted.

However, a closer look at relative clauses involving generalized quantifiers reveals the fact that some languages only allow in these cases the invariant relativizer, corresponding roughly to English that. This is true of a number of Slavic languages including Polish, Serbo-Croatian and Russian. (2) and (3) illustrate this for Polish and SC.

(2) Polish

a. Jan włożył do kieszeni wszystko co/*które mógł.
   Jan put.PAST to pocket all that/which can.PAST
   ‘Jan put in his pocket everything that/*which he could.’

b. Jan przyniósł coś co/*które mnie zaskiczyło.
   Jan brought something that/which me.DAT surprised
   ‘Jan brought something that/which surprised me.’

c. Jan nie przyniósł niczego co/*które mnie zaskoczyło.
   Jan NEG brought nothing that/which me.DAT surprised
   ‘Jan didn’t bring anything that/which surprised me.’

(3) Serbo-Croatian

a. Našao sam sve što/*koje sam želeo.
   found.M.SG AUX.1.SG all that/which AUX.1.SG wanted.M.SG
   ‘I found all I wanted.’

*I am grateful to the audiences at PLC32 and LSRL38 for useful comments and discussion of the material in this paper. Many thanks also to Željko Bošković for detailed discussion of the data and the proposal. The help of my informants is also gratefully acknowledged. All errors and omissions are, of course, my own.

b. Našao sam nešto što/*koje sam želeo.
found.M.SG AUX.1.SG sth. that/which AUX.1.SG wanted.M.SG
‘I found something I wanted.’

c. Nisam našao ništa što/*koje sam želeo.
not-AUX.1.SG found.M.SG nothing that/which AUX.1.SG wanted.M.SG
‘I didn’t find anything I wanted.’

At first sight, this appears to be a property of the Slavic languages. However, at least two counter-examples stand out. Bulgarian, which is also a Slavic language, does not show the restriction at all, and Romanian, a Romance language, shows the restriction, but only with the bare quantifier tot (all/everything). The relevant data is given in (4) and (5), respectively.

(4) Bulgarian
a. Namerih vsichko koeto/deto go iskah.
found.1.SG all which/what it wanted.1.SG
‘I found everything I wanted.’

b. Namerih neshto koeto/deto go iskah.
found.1.SG something which/what it wanted.1.SG

c. Ne namerih nishto koeto/deto go iskah
not found.1.SG nothing which/what it wanted.1.SG

(5) Romanian
a. Ion a pus în buzunar tot ce/*care a putut.
Ion AUX put.PASTP in pocket all that/which AUX can.PASTP
‘John put in his pocket everything he could.’

b. Nu am găsit nimic ce/*care să-mi plăcă.
not AUX.1.SG found nothing which SUBJ-me.DAT like.SUBJ
‘I didn’t find anything which I might like.’

c. Am găsit ceva ce/*care mi-a plăcut mult.
aux.1.SG found something which me-DAT-AUX liked.PASTP much
‘I found something which I liked a lot.’

What Bulgarian and Romanian have in common is the fact that they have articles. As I will show in section 4, it is not the article itself but the implications of the availability of D in a language that affect the relativizer options.

In brief, I will argue that this kind of relativizer restriction is triggered by a failure to overtly mark the edge of an intonational phrase, where phases are assumed to determine intonational phrases. In the next section I provide evidence for the role of intonational phrasing.

3 The Role of Intonational Phrasing

3.1 Evidence from Romanian

An interesting fact about the quantifier tot in Romanian is its ability to combine with nouns bearing the definite article or to bear the article itself:

(6) Tot zaharul/Totul e pe jos.
all sugar.the/all.the is on floor
‘All the sugar/Everything is on the floor.’

1The relativizer ce (‘that,’ literally ‘what’) is generally used in literary language, which makes it a bit unnatural in some informal contexts. Surprisingly, however, it is the only natural option with the bare quantifier tot (‘all’) in the example below.
Surprisingly, however, the presence of the definite article on tot precludes the presence of a relative clause, as in (7), regardless of the chosen relativizer, which is not the case if an overt noun bears the definite article, as in (8):

(7) * Le place totul ce/care depăşeşte limita.
    them.b like all the what/which exceeds limit.the
    ‘They like everything that is beyond the limit.’
(8) Le plac toate lucrurile ce/care depăşesc limita.
    them.b like all things the what/which exceed limit.the

Since the Romanian definite article has to undergo a phonological process (encliticization), I take the above facts as evidence that there is an intonational phrase boundary (IPB) at the level of the DP. IPBs have long been known to prevent the application of certain phonological processes. Many environments have been argued to be obligatorily parsed as separate intonational phrases, e.g., root clauses, parentheticals, tag questions, vocatives, certain moved elements and more recently null-C-disallowing contexts (see Cooper and Paccia-Cooper (1980), Selkirk (1978), Selkirk (1984), Selkirk (1986), Nespor and Vogel (1986), Schütze (1994), Bošković (2001) and An (2007), among others).

I assume that (7) is ungrammatical because the definite article cannot encliticize across an IPB. In order to see how that takes place, let us consider briefly the syntax of relativization. I assume with Kayne (1994) that the relative clause is generated as a complement to D and the head noun is raised from inside the relative clause. The noun moves first to a position at the edge of the relative clause and then to D, via N-to-D movement (see Dimitrova-Vulchanova and Giusti (1998); Bouchard (1998, 2002); Dobrovie-Sorin (2000); Ticio (2003) among others; see also Dimitrova-Vulchanova (2003) for alternative views).

In the case of (7), where no noun is present, encliticization of the definite article on Q needs to take place across an intonational phrase boundary (marked with #), which is responsible for the ungrammaticality. If a noun is present, however, as in (8) above, encliticization does not cross the DP phase/intonational phrase boundary (marked with #); in fact, it is the noun that moves to D.

At this point, one may wonder why DP would determine an intonational phrase. This idea is actually natural given the assumption that intonational phrases correspond to phases, i.e., spellout domains, and the fact that DP has been argued to be a phase (see Bošković (2005) and Svenonius (2004)). However, in order for this account to work, we need to assume that intonational phrases correspond to full phases, not to the recent Chomskyan instantiation of spellout domains (where
only the complement of a phase head is sent to spellout, excluding the specifier).

3.2 Evidence from Serbo-Croatian

In the previous section I have proposed that the inability of the Romanian quantifier tot to be modified by a relative clause when encliticization of the definite article has to apply suggests that intonational phrasing may interfere with this process. In this section I provide further evidence for the existence of an intonational phrase boundary in relative clauses, this time at the level of the CP.

Radanović-Kocić (1988, 1996) and Bošković (2001) have argued convincingly that Serbo-Croatian clitics must occur in the second position of their intonational phrase. Compare the grammatical (11a) and the ungrammatical (11b), which are identical except for the position of the auxiliary clitic sam. Note also that the ungrammaticality is not merely caused by the clitic being preverbal, as evidenced by the grammaticality of (11c).

found.M.SG AUX.1.SG all that AUX.1.SG wanted.M.SG  
‘I found all I wanted.’

b. * Našao sam sve što želeo sam.  
found.M.SG AUX.1.SG all that AUX.1.SG wanted.M.SG  
c. Želeo sam to.  
wanted.M.SG AUX.1.SG it  
‘I wanted it.’

If clitics need to be located in the second position of their intonational phrase, we are forced to conclude that there is an IPB between the quantifier sve and the relativizer što.

Let us now turn to explaining the basic contrast in the paradigm proposed in section 2 above, partially repeated in (12) below.

(12) Našao sam sve što/*koje sam želeo.  
found.M.SG AUX.1.SG all that/which AUX.1.SG wanted.M.SG  
‘I found all I wanted.’

We need to explain why only the invariant relativizer što is allowed in relatives with bare quantifiers. My proposal is based on An (2007), who argues that the edge of an IPB (either the specifier or the head of the phase that is mapped into an IPB) needs to be overtly realized, as shown by the examples in (13). Subject clauses are only grammatical when either the Spec or the head of the CP is overtly filled.

(13) a. [What ∅ John likes] is apples. (Bošković, 1997:182)  
b. [That John likes apples] is widely known.  
c. * [ ∅ John likes apples] is widely known.

Regarding our cases, I propose that the relativizers that are not located in C₀ (Serbo-Croatian koje, Polish ktoře and Romanian care) cause ungrammaticality since the edge of the IPB at the CP-level is not properly marked.

This interpretation of the relativizer restriction has consequences for the syntax of relativization in general. According to Kayne (1994), the difference between the relativizers is only categorial, both being located in the CP, which is incompatible with the present account since it would not predict ungrammaticality in the relevant cases. However, Bianchi (1999) proposes an alternative account, with different positions for the relativizers, shown below for Serbo-Croatian in (14) and (15). The relativizer što, which is located in Force₀ (CP), properly marks the edge of the IPB at the level of the relative clause. On the other hand, koje, which is in Spec,TopP, leaves the ForceP phonetically unrealized, causing ungrammaticality.
4 Explaining the Distribution of the Relativizer Restrictions

What we need to explain is why the Force phrase is sometimes obligatorily parsed as an intonational phrase, which is what I have argued in the previous section is responsible for the relativizer restriction in general (as well as clitic placement in relative clauses). The explanation cannot be merely that all relative clauses are obligatorily parsed as separate intonational phrases. An (2007) argues, based on cross-linguistic data from Tagalog, Brazilian Portuguese, Tuscan Italian and Korean, that restrictive relative clauses may be, but need not be parsed as separate intonational phrases, unlike the closely related noun complement clauses.

With this issue in mind, let us examine the data An (2007) uses to illustrate the contrast between the parsing behavior of noun complements and restrictive relative clauses. According to Richards (1999), Tagalog features both affixal (-ng) and non-affixal (na) complementizers. The examples in (16) and (17) illustrate a difference in the acceptability of these complementizers in noun complements and restrictive relatives. In the noun complement clause in (16b), the affixal complementizer is ungrammatical, unlike in the relative clause in (17). An’s account of these data appeals to a difference in parsing. The affixal complementizer needs to have the preceding nominal as a host, so it is only expected to be grammatical in case there is no IPB in between the affix (-ng) and its host (balita). This is expected if the noun complement clause in (16b) is parsed as a separate intonational
phrase, which makes the affixal complementizer unacceptable; in the relative clause in (17), the affixal complementizer is acceptable, which suggests that the relative clause need not be parsed as an intonational phrase.

(16) a. ang balita [na kinain ni Juan ang tambakol] (NC)
   news that ate Juan mackerel
   ‘the news that Juan ate the mackerel’

b. * ang balita [-ng kinain ni Juan ang tambakol] (NC)
   news that ate Juan mackerel (Richards, 1999)

(17) ang balita [-ng dinala ni Juan] (RC)
   news that brought Juan
   ‘the news that Juan brought’ (Richards, 1999)

The same distinction can be made with respect to English. The noun complement clause in (18a) cannot be introduced by a null complementizer—a null affix, as proposed by Pesetsky (1992) (see also Ormazabal (1995) and Bošković and Lasnik (2003) for further discussion). This is not the case in the corresponding restrictive relative in (18b).

(18) a. The claim that/*∅ Mary offended Bill is unsubstantiated. (NC)

b. The claim that/∅ the defense rested on was supported by the witness. (RC)

The data above are consistent with a view along the lines of Stowell (1981) who treats noun complements as appositives, which are always obligatorily parsed as separate intonational phrases. Relative clauses, on the other hand, are only optionally parsed as intonational phrases, as evidenced by the grammaticality of the null complementizer in (18b). I propose that in the case of (18b), the noun above the CP plays a relevant role in determining how the relative clause is parsed. The top layer of the relative clause can optionally be parsed together with the preceding noun. According to Bianchi (1999), the null C option for relative clauses (and declarative sentences) is really a reflection of the noun selecting for the phonetically null head Topic\(^0\). The resulting structure is shown in (19).\(^2\) If the IPB were located at the level of TopP, (18b) would be expected to be ungrammatical, contrary to fact. If the TopP layer is parsed together with the noun in Spec,AgrP, no ungrammaticality is expected. This also explains why relative clauses with overt nouns, such as those in Tagalog above, are only optionally parsed as separate intonational phrases.

\(^2\)In the case of sentences with an overt noun, which under Bianchi’s analysis described in the text above is placed in Spec,ForceP, the IPB would have to fall between the specifier and the head of the ForceP (in order to have the clitic in the second position of the intonational phrase in (i)(a)), an undesirable consequence. It is therefore necessary to adopt a slightly different version of the above structures, also suggested by Bianchi (1999:200), which features the head noun of the relative clause in an AgrP projection above the ForceP.

(i) (a) svi ljudi # što su došli
   all men that AUX.CL come
   ‘all men who came’

(b) * svi ljudi # što došli su
   all men that come AUX.CL

This allows us to maintain the idea that the intonational phrase boundary is at the level of the ForceP.
4.1 Relativizer Restrictions in Languages without Articles

Let us now return to the issue of the relativizer restriction in Serbo-Croatian introduced in section 2 and discussed in section 3. In (14–15) there is an intonational phrase boundary at the ForceP level, which triggers the relativizer restriction, as a result of which što, but not koje, is possible in the context in question.

However, Serbo-Croatian relative clauses with an overt noun do not exhibit this restriction. Interestingly, also, relatives with the bare quantifier sve show an improved ability to combine with koje (which) when the quantifier agrees with a noun previously mentioned in the context:

(20) Context: There are 15 girls in the choir, but 5 of them are home sick. Only 10 showed up for the show.


In order to understand why that may be the case, it may be helpful to consider relative clauses with overt relativized nouns, which do not evince any relativizer restrictions, as seen in (22) below.

(22) Sve devojke što/koje su došle pevale su all.FEM.PL girls which.FEM.PL AUX come.3.FEM.PL sung.3.FEM.PL AUX punim srcem. full.INSTR heart.INSTR

The absence of the relativizer restriction may be at first attributed to the presence of the overt noun devojke in Spec,CP, in line with Bianchi’s original proposal, which means that the IPB needn’t be marked by the relativizer, making koje acceptable.

However, recall from the discussion in section 3 that Serbo-Croatian clitics are required to be in the second position in their intonational phrase. Given this, a question arises regarding the grammaticality of (22) with koje, whose structure is shown in (23). An IPB at the level of CP in (23) would not be properly phonetically marked, and would therefore trigger ungrammaticality. Moreover, optionally parsing the CP as an intonational phrase does not yield the right results since the clitic would be located in the fourth position in the whole relative clause.
I propose that the optionality of parsing the relative clause as a separate intonational phrase, discussed for English above, should not be interpreted to mean that the noun and the whole relative clause form an intonational phrase together, but rather that the CP/ForceP layer can be optionally parsed with the noun instead of with the rest of the relative clause. Consider again (22) when the relativizer *koje* is chosen. The two parsing options are as follows:

(24) a. the CP is parsed as a separate intonational phrase: the clitic *su* is in 2nd position, but IPB is not properly marked; leads to ungrammaticality
b. the noun *devojke* and the rest of the CP/ForceP layer are parsed separately; the relative clause starting with the TopP layer is parsed separately: the clitic *su* is in 2nd position; leads to grammaticality

When *štto* is selected, the options are the same, but it is the first option that results in grammaticality:

(25) a. the CP is parsed as a separate intonational phrase: the clitic *su* is in 2nd position and the IPB is properly marked by the presence of *štto*; results in grammaticality
b. the noun *devojke* and the rest of the CP/ForceP layer are parsed separately; the relative clause starting with the TopP layer is parsed separately: the clitic *su* is in 1st position; leads to ungrammaticality

We can now try to explain the difference between relatives with *sve* without a noun and relatives with *sve* which agree with a contextually specified noun. I propose that Serbo-Croatian has two *sve* elements. One is invariant and never appears with a noun, overt or inferred; the other always agrees with a noun, either overt or covert. I argue that only the invariant *sve* triggers the relativizer restriction, since no noun is available for the CP layer to be parsed together with. On the other hand, when the agreeing *sve* is used, a noun is always available at the point when intonational phrasing takes place. In the cases where agreeing *sve* is not followed by an overt noun, the noun has undergone PF deletion following intonational phrasing.

The remaining issue is the behavior of the Serbo-Croatian generalized quantifiers *neštto* and *ništta*, which also trigger a relativizer restriction, as seen in (3b) and (3c) in section 2. These quantifiers are also invariant, so I will assume that they behave essentially like invariant *sve*. I will return to the relevance of the nominal structure, i.e., the availability of D, when I discuss Romanian below.

4.2 Relativizer Restrictions in Languages with Articles/D

Let us now turn to accounting for the relativizer restrictions in languages with articles, such as Bulgarian and Romanian. Recall from the previous discussion that Bulgarian shows no relativizer
restrictions, while Romanian only shows a relativizer restriction in one case: when the universal quantifier *tot* is used unaccompanied by a noun or a definite article.\(^3\)

With respect to the universal quantifier *tot*, Romanian behaves just like Serbo-Croatian, discussed in the previous section. We can therefore assume that Romanian also has two quantifiers *tot*: an invariant one, which triggers a relativizer restriction, and an agreeing one, which is always accompanied by a covert or overt noun. Recall that it is the presence of this noun that is responsible for the optionality in the location of the IPB of the relative clause, as seen for Serbo-Croatian above. The CP layer has the option of being parsed with the noun, to the exclusion of the material in TopP and below.

On the other hand, the quantifiers *ceva* (something) and *nimic* (nothing) do not trigger a relativizer restriction, unlike in Serbo-Croatian, a fact which needs to be accounted for. Essentially, relatives involving these quantifiers behave just like those where a noun is projected. It is therefore possible that these quantifiers are the phonetic realization of a nominal projection incorporated into D (cf. Bianchi’s AgrP), since they cannot combine with nouns, but only with adjectives, as shown in (26).

\[(26) \quad \text{ceva} \quad \text{bun/*creion} \quad \text{something good/*pencil} \]

In fact, the English counterparts of *ceva* and *nimic* do overtly what Romanian may do covertly, i.e., they incorporate a noun: *thing*. The Romanian *ceva* is also likely bimorphemic, though less transparently so. *Va* can combine with all the wh-forms, including *ce* (what), yielding indefinite quantifiers: *cineva* (someone), *careva* (anyone), *undeva* (somewhere), etc. Here the *va* morpheme corresponds to a nominal while *ce*- is located in the DP, which also straightforwardly accounts for the data in (26), both the impossibility of combining with a noun and the word order with adjectives. Given what I have said above about the parsing effect of an overt nominal in the relative clause, I assume that the nominal part of these quantifiers is responsible for optionality in the location of the intonational phrase boundary in this case as well.

Before closing this section we need to discuss the case of Bulgarian, which by now is expected to behave like Romanian since it has a definite article, a rare feature among the Slavic languages. Unlike in Romanian, the bare quantifier *vsičko* always allows both relativizers: the invariant *deto* and the agreeing *koeto*. In light of our analysis of Romanian and Serbo-Croatian one possible explanation is that non-agreeing *vsičko* is simply absent in Bulgarian. For the other generalized quantifiers (*nešto* and *ništo*), the morphological analysis presented for Romanian carries over straightforwardly.

5 Summary and Conclusions

In this paper I have presented an account of novel data involving the distribution of relativizers in some Slavic languages and in Romanian. I have argued that the observed relativizer restrictions arise as a result of the failure to satisfy the conditions for marking an intonational phrase boundary at the level of the relative clause. This approach has consequences for the syntax of relativization. It leads to the conclusion that the two kinds of relativizers (invariant and agreeing) are located in different phrases at the left periphery of the clause (ForceP and TopP, respectively). Moreover, I show that the structure of the nominal domain, in particular the availability of D, plays a role in determining the relativizer options.

References


\(^3\)See the data in (7), which shows that *tot* accompanied by a definite article cannot be modified by a relative clause. I have already given an account of these facts, which do not merely illustrate a relativizer restriction, since all relativizers are bad.


