



UDC 811.163.42'367.626:811(4)=111

Received on 14. 07. 2015

Accepted for publication on 08. 02. 2016

Original scientific article

Jurica Polančec
Ana Mihaljević
University of Zagreb

Croatian non-standard invariant relativizers in European perspective

In this paper, the authors analyze Croatian invariant relativizers (InvRels) in European perspective, with emphasis on InvRels employed in non-standard Croatian. The paper is roughly divided in two parts. In the first part of the paper, previously researched facts about InvRels are reviewed. It is demonstrated that in Croatian, InvRels are more typical of non-standard variety and with fewer syntactic restrictions when compared to standard Croatian. Sociolinguistic variation is also briefly commented on. In the second part of the paper, two previously unanalyzed features of non-standard InvRels are examined by means of the *hrWac 2.0* corpus: distribution across the position relativized, and the use of resumptive pronouns with direct object. Due to the nature of corpus data, the investigation was narrowed to the InvRels *šta* and *kaj*. The paper demonstrates that the two InvRels are found in most cases with subject, and to a lesser extent with direct object relativization. Other positions relativized are rarely encountered. Resumptive pronouns with inanimate direct objects are readily omitted, while with animate direct object the omission seems fairly rare. These findings are important because they display that only subject and direct object are regularly relativized, even in non-standard Croatian. Furthermore, the major role of animacy with respect to the use of resumptive pronouns is pointed out, with a comparable tendency in Czech (Fried 2010).

Key words: relative clauses; non-standard varieties; Croatian language; invariant relativizer; corpus-based analysis.



1. Introductory remarks¹

In a number of European languages, regardless of their native speaker numbers, most attention has been devoted to their respective standard varieties. Many of these languages have long traditions of codified standard varieties. Yet, recent times have seen an increased interest for non-standard varieties² of many European languages. Much of that interest has focused on grammatical peculiarities with regard to standard varieties, as well as the sociolinguistic status of non-standard varieties in individual languages. One of the major facets of non-standard varieties in many European languages is the use of relativization strategies distinct from those employed in standard varieties. As several authors pointed out (Cristofaro and Ramat 2007; Fiorentino 2007), standard European varieties have a tendency to employ relative pronouns, such as French *qui/que* or Croatian *koji*, while non-standard varieties more frequently resort to invariant relativizers (henceforth InvRel).³ These observations were substantiated by Murelli (2011), who in his work offers a broad typological investigation of relativization strategies in European languages, its major contribution to the field being the focus on non-standard varieties. Moreover, Murelli (2011) effectively confirmed the aforementioned claims by demonstrating on a broader sample that, among relativization strategies, the relative pronoun strategy and the strategy with InvRels are indeed predominant in most of Europe,⁴

¹ The following abbreviations are used in this paper: ACC ‘accusative’, AH ‘Accessibility Hierarchy’, ANIM ‘animate’, AUX ‘auxiliary’, DO ‘direct object’, F ‘feminine’, GEN ‘(possessive) genitive’, INAN ‘inanimate’, INS ‘instrumental’, InvRel ‘invariant relativizer’, IO ‘indirect object’, M ‘masculine’, NEG ‘negation’, NOM ‘nominative’, OBL ‘oblique’, OCOMP ‘object of comparison’, PRS ‘present’, PTCP ‘participle’, Q ‘question particle’, RC ‘relative clause’, RP ‘relative pronoun’, SG/PL ‘singular/plural’, SUB ‘subject’.

² By non-standard, we mean all varieties subsumed under labels such as “non-acknowledged, non-prestigious, non-codified, non-elaborated” (Murelli 2011: 32). This notion also covers all constructions that “codifiers usually mark as spoken, colloquial, dialectal, rural, etc.” (Murelli 2011: 33). A more specific characterization of what is specifically meant by ‘non-standard’ in Croatian will be given in 3.2.

³ Invariant relativizers are defined as relative elements that encode only the link between the main and the relative clause (Murelli 2011: 123). ‘Relative particle’ is another frequently used term (e.g. by Murelli 2011). InvRels are notoriously difficult to determine with respect to their categorial status. This is still a much disputed topic in literature and it will not be pursued here. For that particular reason, we have opted for the neutral term ‘invariant relativizer’.

⁴ Other relativization strategies reported in Murelli (2011) are the strategy with so-called specialized relative elements, the zero-marker strategy, as well as various instances of strategies making use of combined relative elements (other than the combination of an InvRel and a resumptive pronoun). These strategies are, however, attested to a much lesser extent (see Murelli 2011: 97 and 99).



the former being more typical of standard, and the latter of non-standard varieties. He also showed that this was somewhat oversimplified since, in most languages, InvRels are also found in standard varieties, but they are subject to numerous restrictions there.

These restrictions mostly have to do with the Accessibility Hierarchy (AH), which can be posited as follows (Keenan & Comrie 1977: 66):

SUB > DO > IO > OBL > GEN > OCOMP

Here, > means ‘is more accessible than’. An important claim within the AH is that “any RC-forming strategy must apply to a continuous segment of the AH” (Keenan and Comrie 1977: 67), that is, if a language can relativize OBL, it can also relativize all higher positions, in this case IO, DO, and SUB. In terms of the AH, relative pronouns can be used in European languages to relativize all positions. Unlike them, InvRels in European languages are much less frequently found to relativize positions lower than DO. When they do, they necessitate the use of resumptive pronouns with the InvRel. These restrictions, however, tend to be more relaxed in non-standard varieties.

The facts presented above outline some major tendencies in European languages. Individual languages, on the other hand, fit into the overall picture to a varying degree. Croatian can serve as an illustration of this variation. Specifically, it was demonstrated in Polančec and Gnjatović (2014) that, as is the case in most other European languages, Croatian displays differences in RC formation between standard and non-standard varieties. These differences follow the previously observed pattern: while the relative pronoun and the InvRel are both attested in standard variety, the former largely predominates, while in non-standard varieties the use of InvRels is much more common and subject to fewer restrictions than is the case in the standard variety. Due to a lack of previous studies on InvRels in non-standard Croatian, Polančec and Gnjatović (2014) was mostly heuristically oriented in the sense that it sought to demonstrate a widespread presence of such constructions, and, additionally, to establish the important features of their usage. In particular, it aimed to compare features of non-standard InvRels to the already described features of the InvRel in standard Croatian,⁵ as well as to the features of the predominant relative strategy, the one making use of the relative pronoun *koji*.

⁵ As it will become obvious in Section 3, standard Croatian has a single InvRel, namely *što*, whereas in non-standard, there are four InvRels (*što*, *šta*, *kaj* and *ča*). That is why we will always talk of the InvRel in standard Croatian, and of InvRels in non-standard Croatian. When no standard or non-standard varieties are specifically meant, we will refer to them as InvRels in Croatian in general.



This paper is an attempt to further contribute to the research of InvRels in non-standard Croatian. Its aims are twofold. First, it will use the findings and conclusions in Polančec and Gnjatović (2014) to compare the features and the sociolinguistic status of InvRels in Croatian to findings from other European languages, as presented in Murelli (2011). This is necessary as no analysis of the position of Croatian InvRel constructions in typological and/or areal perspective was offered in Polančec and Gnjatović (2014). Second, it will contribute to the investigation of formal features of InvRels in non-standard Croatian by using quantitative data from language corpora. More specifically, it will investigate the frequency of relativization of lower positions on the AH and the use of resumptive pronouns with DO-relativization by means of the *hrWac* 2.0 internet corpus. For the purposes of this investigation, two InvRels found exclusively in non-standard Croatian, *šta* and *kaj* were chosen. The choice of topic and the rationale for choosing these two InvRels will be laid out in Sections 3 and 4. Data obtained in this fashion and its implications will be analyzed in broader typological terms as well.

It is worth noting that most of the data collected by Murelli came from grammars of European languages, which offer very little information on RC formations in non-standard varieties (Murelli 2011: 57–60). Murelli used language-specific in-depth studies only to a limited extent, mostly due to the fact that they are lacking for many European languages. Bearing this in mind, this paper will attempt to remedy this lack of in-depth studies, principally by presenting established facts about Croatian non-standard InvRels so far and, also by offering some fresh empirical evidence. Our conclusions thus may prove interesting for the considerations of RCs in a Slavic, but also in a broader European perspective.

2. Invariant relativizers in European languages

Two major relativization strategies in European languages, as announced in the introduction, are presented here: the relative pronoun strategy, and the strategy with InvRels. Their major features, geographical distributions, and sociolinguistic statuses are reviewed. This overview is largely based on Murelli (2011). It serves as a basis for the comparison of Croatian data to other European languages in Section 3.

First, the relative pronoun strategy is illustrated with examples from English in (1a) and Czech in (1b):⁶

⁶ In order to facilitate the reading, in every example the head will be underlined and the RC will be closed in square brackets. The relativizers will be in bold, as will the resumptive pronouns when used together with an InvRel.



- (1) a. RC with the relative pronoun *who* from English (Quirk et al. 1985: 1244):

They are delighted with the person [who has been appointed].

- b. RC with the relative pronoun *který* from Czech (Fried 2010: 1):

| | | | | | |
|------------------|-----------|-------------------|------------------|----------------|-------------|
| <i>Vyplývá</i> | <i>to</i> | <i>z</i> | <i>dohody,</i> | <i>[kterou</i> | <i>jsme</i> |
| follow.3SG | it | from | agreement.GEN.SG | which.ACC.SG.F | AUX.1PL |
| <i>podepsali</i> | <i>s</i> | <i>vedením</i> | <i>firmy]</i> | | |
| sign.PTCP.PL | with | management.INS.SG | company.GEN.SG | | |

‘It follows from the agreement [that] we signed with the company’s management.’

In the relative pronoun strategy, “the position relativized is indicated inside the relative clause by means of a clause-initial pronominal element, and this pronominal element is case-marked (by case or by an adposition) to indicate the role of the head noun within the relative clause” (Comrie and Kuteva 2013a). As for our examples, in (1a) the head noun *person* occupies the subject position⁷ in the relative clause, which is signaled by the relative pronoun *who*, whereas in (1b) the head noun *dohody* ‘agreement’ occupies the direct object position in the RC, which is in turn signaled by the accusative case of the relative pronoun.

We now turn to InvRels, which are again illustrated with examples from English (2a) and Czech (2b):

- (2) a. RC with the InvRel *that* from English (Biber et al. 1999: 615)

Then the woman [that they actually caught and pinned down] would not have been Margot.

- b. RC with the InvRel *co* from Czech (Fried 2010: 2)

| | | | | |
|------------|------------|-----------|----------------------------|---------------|
| <i>zda</i> | <i>na</i> | <i>mě</i> | <i>čeká</i> | <i>holka,</i> |
| whether | for | me | wait.PRS.3SG | girl.NOM.SG |
| <i>[co</i> | <i>jsm</i> | <i>ji</i> | <i>{měl nadevše rád}</i> | |
| InvRel | AUX.1SG | she.ACC | {loved more than anything} | |

‘whether the girl, who I loved more than anything, is waiting for me’

⁷ In this paper we will use the convention that when the syntactic position is mentioned, e.g. SUB, it always refers to the position relativized.



The example (2a) from English makes use of the InvRel *that*, which alone is sufficient to convey the information that the head noun *woman* has the role of DO in the RC. Since the head is left unrepresented in this way in the RC, this relativization strategy is called the gap strategy. Contrary to English, in Czech the InvRel *co* alone is not sufficient to signal the role of the head in the RC. To this end, the resumptive pronoun *ji* is employed. Standing in the accusative case, it indicates that the role of the head in the RC is that of DO. Thus, the resumptive pronoun represents the head in the RC, and therefore this kind of relativization strategy is called the pronoun-retention strategy. Now, let us look into the distribution of the the relative pronoun strategy and the two strategies employed with InvRels in European languages.

The relative pronoun strategy is commonplace in European languages: in Murelli's sample, it is attested in 32 out of 36 languages (2011: 97, 244).⁸ Apart from its widespread use in Europe, this relativization strategy is virtually non-existent in non-European languages (Comrie and Kuteva 2013a, 2013b). Its frequency in and conspicuous scarcity outside of Europe makes it one of the prominent features of the European *Sprachbund* (Haspelmath 2001).⁹ As for RC formation in non-European languages, several other strategies are attested (e.g. Comrie 1989: 138–164). Two of them have already been mentioned above: the gap and the pronoun-retention strategy. The former is largely predominant in Southeast Asia, the Pacific area, and Australia, and is also frequently found in Sub-Saharan Africa. Other strategies attested outside Europe are internally headed relative clauses, most frequently employed in the Americas, and correlative clauses, which predominate in South Asia.

⁸ Interestingly enough, languages lacking the relative pronoun strategy are languages on the European periphery: Icelandic, Irish, Maltese and Turkish.

⁹ The exact causes and/or mechanisms which led to the rise of the relative pronoun strategy in Europe are still not well understood. The RCs featuring relative pronouns may have arisen owing to the influence of Latin and other prestigious languages in the Middle Ages (cf. Haspelmath 2001: 1494–1495, Fiorentino 2007: 282–284). Following that scenario, RCs featuring InvRels can be seen as original in European languages, and it can be argued that the prominence of InvRels in written languages must have decreased due to the spread of relative clauses in imitation of more prestigious varieties. The evidence from the languages that still feature InvRels as their predominant strategy also points to this direction (cf. Murelli 2011: 270), as they are again mostly found on the European periphery (cf. note 8), where in particular the influence of Latin was not as strong (e.g. Scandinavian languages, Irish, Maltese, and Icelandic). See also Murelli (2011: Ch 6), and for the notion of prestigious language Murelli (2011: 257).



Let us now turn to InvRels. In Murelli's sample, InvRels used alone, i.e. instances of the gap strategy, are in fact attested in more languages than relative pronouns, namely in 34 out of 36.¹⁰ Instances of the pronoun-retention strategy, i.e. InvRels used with resumptive pronouns, are attested in 31 languages of his sample. In the introduction, it was pointed out that InvRels are normally found in non-standard varieties of European languages. This is by no means an absolute rule, as there are languages where InvRels are not confined to non-standard varieties. According to Murelli (2011: 265–268), four groups of languages can be outlined as to the constraints on the use of InvRels in their respective standard varieties. In languages belonging to Group 1, InvRels are not codified in the standard. In Group 2, InvRels are codified in the standard for SUB and DO-relativization only. Croatian is listed among these languages. In Group 3, the use of InvRels in standard language is not restricted to SUB and DO-relativization, but the use of resumptive pronouns is obligatory for all positions except SUB and DO.¹¹ Finally, in Group 4, comprising of only one language, Greek, the InvRel with no resumptive pronoun can be used for all positions in standard language.¹²

The reported findings merit further discussion, since the question poses itself why InvRels are dispreferred to relative pronouns in European standard varieties. We have seen that InvRels alone are acceptable in most standard varieties as far as SUB and DO-relativizations are concerned. By contrast, relativization of lower positions, when codified in the standard, necessitates the use of resumptives, Greek being the only exception. Drawing on the existing literature on standardization principles in European languages (e.g. Stein 1997; Ammon 2004), Murelli (2011: 259ff) argues that these observations can be accounted for by invoking two features of European standard varieties, compactness and explicitness.¹³ The factor of com-

¹⁰ The only two languages with no InvRels attested are Estonian and Hungarian.

¹¹ Group 3 can be broken down into several sub-groups. In the sub-group consisting of four Slavic languages (Macedonian, Slovak, Slovenian and Ukrainian), InvRels are used together with relative pronouns. In the sub-group consisting of Scandinavian languages and English, relative pronouns are attested, but InvRels are more common. This is particularly true of Scandinavian languages, where relative pronouns are only found in very formal language. In English, relative pronouns are more common than in Scandinavian languages, and the InvRel *that* is shown to be somewhat more frequent (Biber et al. 1999: 610–611). In the remaining three languages of Group 3, Icelandic, Irish and Maltese, no relative pronouns are found, and in these languages, InvRels are the predominant strategy. In connection to these three languages, cf. notes 8 and 9 above.

¹² There are ten other languages attested where InvRel with no resumptive pronoun can be used for all positions, but this use is restricted to non-standard varieties (Murelli 2011: 276).

¹³ According to Murelli, compactness “has to do with economy of means: if an analytic and a synthetic form exist, the synthetic form is preferred; if two different forms are attested, one of which



pactness disfavors all uses of InvRels employing resumptives, since relative pronouns encode both the link between the main and relative clause and the position relativized, i.e. they condense more morphosyntactic meanings. On the other hand, the factor of explicitness disfavors the use of InvRels alone for positions lower than DO, since InvRels alone cannot signal the position relativized and since, in such cases, the task of recovering the role of the head in the RC is left to the hearer.

To summarize, relative pronouns and InvRels (alone or with resumptives) are shown to be the most commonly found relativization strategies in European languages. Still, the two strategies diverge considering their sociolinguistic status, in a way that the presence of InvRels is reduced in standard varieties. As presented above, if InvRels are attested in standard varieties, they are predominately used with no resumptives to relativize only higher positions of the AH, i.e. SUB and DO. Conversely, in non-standard varieties, InvRels are far more frequently used with resumptives or alone to relativize lower positions of the AH.

A final remark is in order. Relativization strategies in some languages, when investigated in an in-depth corpus-driven study, seem to reveal a far more complex picture than presented in Murelli. This can be seen in Czech. In Murelli's classification presented above, Czech is listed in Group 2, that is, as a language where no resumptive is used for DO-relativization. However, we saw that in the example (2b), cited from Fried (2010), the resumptive pronoun *ji* was employed with DO. Moreover, Fried (2010) reports that in Czech, resumptive pronouns are optional with DO-relativization only when DO is inanimate.¹⁴

A similarly complex situation is found in Croatian, also classified as part of Group 2 by Murelli. Presenting this complex situation (Section 3) and contributing to the understanding of some phenomena related to it (Section 4) is going to be the topic of the remainder of the paper.

condenses more distinct (morphosyntactic) meanings in itself, this form is preferred" (Murelli 2011: 260), whereas explicitness "implies that each meaning (i.e. each piece of semantic, morphosyntactic and/or pragmatic information) needs to be encoded through a linguistic form, in order to avoid any kind of direct reference to actual speech situation or relying on pragmatic inference" (2011: 260–261).

¹⁴ Among other corpus-driven studies, we can mention Weinert (2004), for German, and Sonnenhauser (2013), for Slovenian.



3. The invariant relativizer in Croatian

3.1. *InvRel and the relative pronoun koji*

In this section, we are going to present the rules governing the use of InvRels in Croatian. In Croatian, as in most European languages, the predominant relativization strategy is the one making use of the relative pronoun *koji* ‘who, which’,¹⁵ regardless of the register. This is illustrated in (3a) and (3b):¹⁶

- (3) a. RC with the relative pronoun *koji*

| | | | | |
|--------------|-----------|-----------------|---------------|-----------------|
| <i>To</i> | <i>su</i> | <i>putovi</i> | [<i>koji</i> | <i>povezuju</i> |
| this | are | path.NOM.PL | RP.NOM.PL.M | connect.PRS.3PL |
| <i>Istok</i> | <i>i</i> | <i>Zapad</i>]. | | |
| East.ACC.SG | and | West.ACC.SG | | |

‘These are the paths which connect East and West.’

- b. RC with the relative pronoun *koji*

| | | | | |
|--------------------|------------|--------------------|---------------|------------|
| <i>Prodali</i> | <i>smo</i> | <i>kuću</i> | [<i>koju</i> | <i>smo</i> |
| sell.PTCP.PL | AUX.1PL | house.ACC.SG | RP.ACC.SG.F | AUX.1PL |
| <i>naslijedili</i> | <i>od</i> | <i>djeda</i>]. | | |
| inherit.PTCP.PL | from | grandfather.GEN.SG | | |

‘We sold the house that we had inherited from our grandfather.’

Under certain circumstances the relative pronoun *koji* can be replaced in standard Croatian by the InvRel *što*. We illustrate the InvRel *što* with (4a) and (4b), which repeat (3a) and (3b), respectively, with the pronoun *koji* replaced by the InvRel *što* in both examples.

- (4) a. RC with an InvRel

| | | | | |
|--------------|-----------|-----------------|--------------|-----------------|
| <i>To</i> | <i>su</i> | <i>putovi</i> | [<i>što</i> | <i>povezuju</i> |
| This | are | path.NOM.PL | InvRel | connect.PRS.3PL |
| <i>Istok</i> | <i>i</i> | <i>Zapad</i>]. | | |
| East.ACC.SG | and | West.ACC.SG | | |

¹⁵ There are several other relative pronouns in Croatian. For a cursory overview see Polančec and Gnjatović (2014: 193–195) and the references listed there.

¹⁶ Both examples are constructed.



‘These are the paths which connect East and West.’

b. RC with an InvRel

| | | | | |
|----------------|--------------------|--------------|--------------------|------------|
| <i>Prodali</i> | <i>smo</i> | <i>kuću</i> | [<i>što</i> | <i>smo</i> |
| sell.PTCP.PL | AUX.1PL | house.ACC.SG | InvRel | AUX.1PL |
| <i>je</i> | <i>naslijedili</i> | <i>od</i> | <i>djeda</i>]. | |
| she.ACC | inherit.PTCP.PL | from | grandfather.GEN.SG | |

‘We sold the house that we had inherited from our grandfather.’

The example (4a), parallel to (3a), makes use of the InvRel alone, as the InvRel is enough to signal that the head noun occupies the SUB position in the RC. On the other hand, in (4b), parallel to (3b), the role of the DO assumed by the head noun in the RC is signaled by the resumptive pronoun *je*.¹⁷

In Croatian, along with the InvRel *što*, there are three other forms, viz. *šta*, *kaj* and *ča*. However, in standard Croatian, the only form accepted is *što*. In non-standard Croatian, all four InvRels (*što*, *šta*, *kaj* and *ča*) are employed. For more on these various forms, see Section 3.2. Meanwhile, in the remainder of the section, we will address the differences between the relative pronoun strategy and the InvRel strategy. The InvRel *što* used in standard Croatian will be presented first, followed by the features specific to the non-standard InvRels.

The main difference between the strategy with *koji* and the strategy with the InvRel *što* in standard Croatian relates to the Accessibility Hierarchy, i.e. to the question of which positions in the RC are accessible to relativization.¹⁸ In Standard Croatian, the relative pronoun *koji* has no constraints in this respect, whereas the InvRel *što* can relativize neither GEN nor OCOMP positions (Kordić 1995: 157–158).¹⁹ As for OBL relativization, opinions differ. Some claim that the InvRel *što* in

¹⁷ In Croatian, the 3rd person personal pronouns are employed as resumptives.

¹⁸ There are some problems with using the label OBL in Croatian, since Croatian makes use of various cases and prepositions to indicate both arguments and adjuncts. In that respect, it is not clear whether there should be more positions indicated in the AH for Croatian. Nevertheless, we will use the existing AH by assuming that the IO position includes prepositionless dative and genitive cases occupying argument positions, while OBL covers all PPs and the instrumental case. The label PREPO is probably used along a similar line instead of OBL by van der Auwera and Kučanda (1985: 925).

¹⁹ Van der Auwera and Kučanda (1985: 925) claim that there are no differences with respect to this criterion, even though they admit that the OCOMP is only marginally accessible, which entails that the GEN is fully accessible to *što*. However, that claim cannot be maintained, as our data presented in 4.1 will further demonstrate.



standard Croatian can be used to relativize OBL in various functions (e.g. Silić and Pranjković 2005: 355), and others that it cannot (Kordić 1995: 157). It can be nonetheless noted that such examples are practically non-existent in standard Croatian since they would sound too cumbersome for its stylistic preferences.²⁰

Another important aspect of the description of InvRels is the use of resumptive pronouns. Resumptive pronouns, as mentioned before, are used to indicate the syntactic role of the head in the RC, something an InvRel alone cannot. In standard Croatian, the use of resumptive pronouns is not possible with SUB, but is obligatory for all other accessible positions, i.e. for DO, as in (4b) above, and also IO, where dative or genitive pronouns are used. The final major difference between the two strategies seems to include the tendency to perceive the use of *što* as incorrect or at least less acceptable when the identity of the referent serving as the head of the RC is not yet established (Browne 1986: 81, cf. Polančec and Gnjatović 2014: 198).

Apart from the grammatical rules mentioned above, the InvRel *što*, interestingly, is not equally acceptable in all registers of standard Croatian. The InvRel *što* is more typical of very formal, more elevated registers, especially in the language of literature (Pranjković 1986). In other registers of standard Croatian, such as journalistic or scientific, the InvRel *što* is much less used. When it is used, the InvRel *što* is often a mere substitute for the relative pronoun (Kordić 1995: 164).

Compared to the typology presented in the previous section, according to which most standard varieties allow for SUB and DO-relativization with InvRels, standard Croatian seems to fit in well. The only difference is that, in standard Croatian, IO-relativization is also readily found, which probably stems from the fact that resumptive pronouns for both DO and IO are clitics. Under that assumption, OBL-relativization is dispreferred or even ruled out due to the fact that OBLs, which are always preceded by a preposition, can never be clitics.²¹ Accordingly, a rule can be posited for standard Croatian stating that the InvRel is allowed, or found appropriate, only when resumptive pronouns are clitics, i.e. only with DO and IO-

²⁰ One of the reviewers found this claim unacceptable due to a lack of a corpus study to confirm it. However, we believe it to be highly unlikely that anyone with the scarcest knowledge of standard language would write down a sentence such as *čovjek što sam gledao u njega* 'lit. *the man that I looked into him*'. As it will be shown in 4.1, such examples are rare even in non-standard Croatian.

²¹ Croatian has the opposition of unstressed (clitic) and stressed personal pronouns for the accusative and dative cases, among other forms. After a preposition, the use of stressed forms of the 3rd person pronouns is mandatory. Cf. also note 17. In addition, the personal pronouns in the instrumental case, which is subsumed under the OBL label (cf. note 18), have no clitic forms, i.e. they are always stressed.



relativization. Additionally, in standard Croatian, resumptives are obligatory with DO-relativization, which is not the case, for instance, in Czech, as presented in Section 2.

We point out again that the aforementioned restrictions on the use of the InvRel *što* are valid for standard Croatian only. More specifically, the use of InvRels in non-standard Croatian and standard Croatian shows some differences, which were first identified in Polančec and Gnjatović (2014). That study was the first to specifically address the issue of InvRels in non-standard Croatian and it was carried out by means of collecting numerous examples of RCs with InvRels from the Internet via the *Google* search engine,²² with some additional examples taken from everyday conversation. The study included all four non-standard InvRels: *što*, *šta*, *kaj* and *ča*. It identified two major syntactic differences between the ways the standard InvRel and its four non-standard counterparts are employed: the first being related to the accessibility of OBL-relativization and the other to the restrictions on the omission of resumptive pronouns.²³

Regarding the accessibility of lower positions on the AH in general, the instances of OBL-relativization with the non-standard InvRels were readily found, though it was impossible to decide on their actual frequency (Polančec and Gnjatović 2014: 206–207). Conversely, parallel to the standard InvRel *što*, neither GEN nor OCOMP-relativization was attested in non-standard Croatian. The issue of positions relativized will be taken up in 4.1, and will include quantitative data.

As for the omission of resumptive pronouns, recall that in standard Croatian, resumptive pronouns are not employed with SUB-relativization, but they are obligatory with DO and IO-relativization. In comparison, Polančec and Gnjatović (2014: 207–208) report that in non-standard Croatian the omission of resumptives is commonly attested with DO-relativization, but solely in cases when the head is inanimate (DO-INAN). As for the instances when the head is animate (DO-ANIM), the omission is scarcely ever attested, leading to the conclusion that in this case, the omission is highly improbable, if not impossible. The issue of omission of resumptives with DO will be covered in more detail in 4.2 below, again including quantitative data.

²² This kind of choice reflected the lack of a Croatian corpus featuring non-standard data at the time when the research was being carried out. The *hrWac* corpus used in the present study became available only later, in April 2014.

²³ Among other salient features, the frequent use of dummy heads and NP-heads preceded by determiners were also reported. These topics will not be covered in this paper.



In terms of comparison with other European non-standard varieties, Croatian non-standard InvRels seem to share the same tendencies: resumptives tend to be omitted; lower positions are more readily relativized.

To summarize, the relative pronoun in Croatian has no restrictions, regardless of the register. InvRels are used in both standard and non-standard Croatian. In standard Croatian, the use of the InvRel *što* is very constrained and mostly confined to more formal registers. Syntactically, the standard InvRel *što* is permitted for SUB, DO and IO-relativization, with resumptive pronouns obligatory for DO and IO. InvRels in non-standard Croatian show fewer constraints: OBL-relativization is more readily found and the omission of resumptive pronouns is frequent with DO-relativization when the head is inanimate.

In the next section, several sociolinguistic issues relevant for Croatian will be commented on. Above all, the four different relativizers mentioned on several previous occasions will be finally addressed.

3.2. *The four non-standard Croatian InvRels*

In the previous section, it was mentioned in passing that the form *što* is the InvRel form used in standard Croatian, whereas in non-standard, there are three other forms along with *što* (*šta*, *kaj* and *ča*). These four forms correspond to interrogative neuter pronouns of the same form, after which the three major groupings of Croatian dialects, Štokavian, Kajkavian, and Čakavian, are named. Thus, the forms *što* and *šta* are used in Štokavian, the form *kaj* in Kajkavian, and the form *ča* in Čakavian dialect. Therefore, we are dealing with four forms that are parallel to each other in the way that they developed in parallel from equivalent sources in their respective dialects, i.e. interrogative neuter pronouns.

This arrangement of InvRels reflects the situation in Croatian dialects. However, dialects, which are mostly confined to rural areas and smaller communities, are not the only component of what we termed ‘non-standard’ variety. Recall that in the introduction, we mentioned that the notion of the non-standard includes all varieties that feature elements not codified in the standard variety. In this particular case, we need to mention the component of non-standard termed ‘everyday colloquial language’, ‘conversational standard’ or ‘informal spoken standard’, which bears much more similarities to standard language than dialects do. This is because this non-standard idiom is in fact a sub-variety of standard language which features some non-codified elements and is often influenced by dialects (van Marle 1997: 13–17; cf. Langston and Peti-Stantić 2014: 30). In the remainder of the paper, we

will use the term ‘(everyday) colloquial Croatian’ when this particular non-standard variety of Croatian is specifically referred to.

Moreover, the four aforementioned InvRels are not equally represented in everyday colloquial Croatian. Since standard Croatian is based on the Štokavian dialect, in colloquial Croatian the forms *šta* and *što* are naturally the most represented of the four InvRels.²⁴ The remaining two InvRels, i.e. *kaj* and *ča*, differ as to their status in colloquial Croatian. Specifically, the form *kaj* is a prominent feature of colloquial Croatian due to the fact that the form *kaj* is widely used in the urban non-standard idiom of Zagreb, Croatia’s capital. This non-standard idiom is nowadays mostly Štokavian with very few Kajkavian elements, and the rather frequent word *kaj* is a remnant of the city’s almost defunct original Kajkavian idiom (Šojat et al. 1998: 61–80).²⁵ Unlike *kaj*, the form *ča* is exclusively confined to islands and smaller communities along the Croatian coast. Its exclusive dialectal status is reflected in its much reduced presence on the Internet, as was already noted in Polančec and Gnjatović (2014), where the examples of *ča* used as an InvRel were rather infrequent when compared to the other three InvRels.²⁶

After this general presentation of Croatian InvRels, in the rest of the paper we turn to the analysis of empirical data on two particular topics, which in turn will help us locate Croatian more precisely in the typology of European RC formation.

4. Analysis of Croatian findings

As mentioned in the introduction, the research conducted in this paper continues the investigation in Polančec and Gnjatović (2014), its focus shifting here to quantitative analysis. The cited study, as specified in Section 3, was mostly carried out by means of collecting examples via the *Google* search engine. Accordingly, that kind of research was almost exclusively heuristic, and no quantitative analysis was possible due to the lack of an established sample. In the meantime, the Croatian

²⁴ Contrary to some claims (e.g. van der Auwera and Kučanda 1985), the forms *šta* and *što* are mutually interchangeable in all uses, the only difference being that *što* is the only form allowed in standard language, whereas in colloquial Croatian both *šta* and *što* are used.

²⁵ The prominent use of the form *kaj* does not mean that Croatian as used today in urban Zagreb is in any way Kajkavian. It is in fact one of the typical instances of colloquial Croatian, and therefore similar to standard Croatian and essentially Štokavian. Because of that, both Štokavian InvRels *što* and *šta*, as well as the InvRel *kaj* are normally used in Zagreb.

²⁶ This is also reflected in the *hrWac* 2.0 corpus: *ča* is 10-times less frequent than *kaj* and 20-times less frequent than *šta*.



web corpus, *hrWac* 2.0 (Ljubešić and Klubička 2014), became available. This corpus is the only Croatian corpus that contains texts with non-standard features, which include forum discussions, blogs as well as commentaries on various Croatian websites. In addition, the interface of the corpus enables us to generate samples which can be analyzed quantitatively.

Additionally, it was announced in the introduction that only the InvRels *šta* and *kaj* will be investigated in this paper. The reasons for narrowing the choice to these two InvRels are based on the facts presented in Section 3.1. Specifically, the form *ča* was excluded due to its low frequency in the *hrWac* 2.0 corpus,²⁷ which is probably the consequence of its largely dialectal status, while the form *što* was excluded on the grounds that it would be impossible to have a sample of examples of *što* pertaining exclusively to non-standard Croatian since this form is used in both standard and non-standard Croatian. Unlike them, the forms *šta* and *kaj* are not confined to a dialect, but they are not found in standard Croatian. These facts led us to the conclusion that for the purposes of the current investigation, the form *šta* and *kaj* were best suited to display the features of non-standard Croatian.

In order to compile two samples, one with *šta*, the other with *kaj*, we confronted a major challenge owing to the fact that the words *šta* and *kaj* have numerous divergent uses,²⁸ so that any kind of sample would need to be disambiguated manually first. Among these different uses, the use as an InvRel is one of the less frequent ones. Bearing that in mind, a sample first needed to be filtered in order to produce another sample where the frequency of an InvRel is increased in a sufficient degree to make manual disambiguation feasible.²⁹ A goal of 500 examples for each InvRel was set. After a preliminary check,³⁰ it was concluded that for a sample of at least

²⁷ See the previous note.

²⁸ For an overview of these uses see van der Auwera and Kučanda (1985: 919–923) and Pranjković (2010).

²⁹ The used filters mostly aimed to exclude combinations of *šta/kaj* and the words that frequently precede them, which in turn unambiguously signal that the words *šta/kaj* are not InvRels. For example, *šta/kaj*, when preceded by a neuter demonstrative pronoun (e.g. *to* ‘that’) can unambiguously be recognized as interrogative pronouns. Likewise, when *šta/kaj* are preceded by some prepositions or adverbs, they form complex conjunctions, e.g. *nakon šta/nakon kaj* ‘after (conj.)’, where *nakon* means ‘after (adv.)’. By removing such combinations, the overall sample was greatly reduced and the manual disambiguation was thus made feasible.

³⁰ The preliminary check was made as follows. A random sample of 100 examples of both *kaj* and *šta* was taken from the reduced samples as described in the previous note. About one in 30 examples with *šta* was an InvRel, and therefore a random sample of 15000 was taken. As for *kaj*, following the same procedure, one in 10 examples was recognized as an InvRel, and therefore a random sample of 5000 was taken.



500 sentences a random sample of 15,000 is needed for *šta* and a random sample of 5,000 for *kaj*. Following manual disambiguation, we arrived at a sample of 1077 sentences for *šta* and a sample of 503 sentences for *kaj*,³¹ which were in turn used in our analysis. The analysis is presented in subsections 4.1 and 4.2.

Before turning to the analysis, two remarks concerning the internet language are in order. First, we need to mention that by term colloquial Croatian we do not mean exclusively Croatian used orally, but rather, as Murelli (2011: 47–52) explains, “conceptionally spoken language”, i.e. language that is written but is conceived as and resembles spoken, typically in internet chats or forum discussions. Second, the language reflected in the *hrWac* corpus features standard and everyday colloquial Croatian. Dialectal varieties, on the other hand, appear to be represented only to a limited extent, which is visible in the reduced presence of the InvRel *ča*, as well as in the reduced presence of dialectal uses of the InvRel *kaj*. As for our sample, we decided to include only the instances of the colloquial use of *kaj*, while excluding scarce dialectal examples, in order to keep the sample representative of everyday colloquial Croatian as defined in Section 3.2.³² With that in mind, the rest of the paper will focus on InvRels as used in everyday colloquial Croatian, and not on non-standard Croatian in general. However, for the sake of simplicity, we will continue to use the term ‘non-standard’ to describe the two investigated InvRels, *šta* and *kaj*.

4.1. Distribution across the position relativized

In this section we present the findings concerning distribution across the positions relativized. First, let us recall that in European non-standard varieties, lower positions of the AH are often more accessible to relativization than in their respective standard varieties. The findings from our sample are presented in Table 1:

³¹ While the estimation was right for *kaj*, *šta* turned out to be two-times more frequent than expected. This led to the discrepancy among the sizes of two samples.

³² In our sample, purely dialectal examples have in fact been quite rare. Instead, most of the examples with *kaj* reflected the colloquial language of Zagreb (cf. Section 3.1 and note 25) and colloquial language employed by speakers of Kajkavian dialect who use the colloquial standard. The latter examples were easily identified as speakers in such examples tend to use the Štokavian colloquial standard with some easily recognizable Kajkavian elements, mostly lexical and sometimes grammatical (e.g. frequent use of contracted forms of the auxiliary *budem*: *bum* or the l-participle ending in *-l* instead of *-o*: *gledal* instead of *gledao*).



Table 1:

| Position relativized | ŠTA | | KAJ | |
|----------------------|-------------|---------|------------|---------|
| SUB | 868 | 80,59% | 361 | 71,77% |
| DO | 181 | 16,81% | 123 | 24,45% |
| Other (IO and OBL) | 28 | 2,60% | 19 | 3,78% |
| Total | 1077 | 100,00% | 503 | 100,00% |

As can be seen from the table, SUB and DO are by far the most frequent positions relativized, with the InvRel *kaj* having a higher share of DO examples. Among other positions, IO and OBL are attested. Their exact numbers are for *šta*: IO 8 examples, OBL 16 examples; and, for *kaj*: IO 9 examples, OBL 10 examples. Therefore, examples such as the following occur in our sample exceedingly rarely:

(5) RC with OBL-relativization in Croatian

Ovi *šta* *sam* *već* *pitao* *za* *njih*
 this.NOM.PL.M InvRel AUX.1SG already ask.PTCP.SG for they.ACC

‘these that I have already asked about’

We found no examples of OCOMP and GEN, which confirms the established view that these positions are not accessible for relativization in Croatian as far as InvRels are concerned (see Section 3.1).

The findings presented here show that one should be cautious when deciding on the prominence of relativization of lower positions on the AH. They show us that, while relativization of a position might be attested in a language, it does not mean that it is regularly used. Therefore, in non-standard Croatian, the only positions regularly relativized are SUB and DO.

4.2. Use of resumptive pronouns with DO

In this section, we will address the issue of the use of the resumptive pronouns with DO-relativization with InvRels *šta* and *kaj*. Recall that in Section 3.1, we have shown that the data on Croatian as represented in Murelli differs from the attested state in the way that, in standard Croatian, the use of resumptives is mandatory for both DO and IO, while in non-standard, and more specifically, colloquial Croatian, there is a certain degree of freedom with DO-INAN. An example with omission from our sample is given in (6):

(6) DO-INAN with omission

Ne znam ni dal tko čita
NEG know.PRS.1SG nor Q anyone read.PRS.3SG

ove gluposti [šta pišem]
this.ACC.PL.F nonsense.ACC.PL InvRel write.PRS.1SG

‘I don’t even know if anyone reads this nonsense that I write’

In this example, the accusative plural form *ih*, which refers to the head noun *gluposti* ‘nonsense’, is omitted. This observation can now be substantiated with quantitative data presented in Table 2:

Table 2:

| DO-INAN | ŠTA | | KAJ | |
|-----------------|--------|---------|--------|---------|
| % in the sample | 13,56% | | 21,47% | |
| None | 101 | 69,18% | 91 | 84,26% |
| Resumptive | 45 | 30,82% | 17 | 15,74% |
| Total | 146 | 100,00% | 108 | 100,00% |

As can be seen from Table 2, omission of resumptive pronouns clearly predominates with both relativizers. The RCs with *kaj* are somewhat freer in this respect, which is interesting bearing in mind the data from 4.1, which shows that *kaj* has a greater share of DO-relativization.

As for DO-ANIM, one may recall that in this case, the use of resumptive pronouns is mandatory, even in non-standard Croatian.³³ Contrary to that claim, several such examples have been encountered in the present sample, and the count can be seen in Table 3:

Table 3:

| DO-ANIM | ŠTA | | KAJ | |
|-------------|-------|---------|-------|---------|
| % in sample | 3,16% | | 2,98% | |
| None | 3 | 8,82% | 6 | 40,00% |
| Resumptive | 31 | 91,18% | 9 | 60,00% |
| Total | 34 | 100,00% | 15 | 100,00% |

³³ In Polančec and Gnjatović (2014: 209), three examples with omission of DO-ANIM resumptives were found. However, these cases were deemed as exceptional and attributed to mistakes in language production. Additionally, the authors’ own intuition strongly suggested ungrammaticality of such sentences.



This is illustrated by the following example with *kaj*:

(7) DO-ANIM with omission

| | | | | | | |
|---------------------------------------|------------|-----------|-------------|-------------|-----------|------------------|
| <i>Jedna</i> | <i>sam</i> | <i>od</i> | <i>onih</i> | <i>[kaj</i> | <i>si</i> | <i>nabrojao]</i> |
| one.NOM.SG.F | am | of | that.GEN.PL | InvRel | AUX.2SG | mention.PTCP.SG |
| 'I'm one of those that you mentioned' | | | | | | |

In this example, the accusative plural form *ih*, which refers to the demonstrative *onih* 'those', is omitted.

With *šta*, examples with omission are quite rare, as expected. With *kaj*, however, there are six such examples, constituting 40% of all examples with DO-ANIM. It cannot be claimed that all of these examples are attributable to errors in language production, although this number of examples does not suffice to establish any rule.

In that light, it stands to reason to conclude that the use of resumptive pronouns with DO-ANIM head is not an absolute syntactic rule, but rather a tendency with a very strong pragmatic and semantic motivation (cf. Fried 2010 with similar conclusions for Czech). This motivation is so strong that examples with omission are by-and-large absent from production as the counts above suggest, but they are possible and sometimes, no matter how rarely, licensed. Finally, it would be interesting to see whether there are other European languages where animacy plays such a decisive role with respect to omission of resumptive pronouns. Currently, we are not aware of any such language besides Croatian and Czech.

5. Conclusions

This paper had two explicit goals. The first was to present Croatian InvRels and to place them in the European context. The other was to broaden our knowledge of the rules governing the use of InvRels in non-standard Croatian with the help of language corpora.

In order to present the known facts about InvRels in Croatian, we have first outlined the findings on relativization strategies in European languages as gathered and analyzed in Murelli (2011). In short, this overview showed InvRels to be more typical of European non-standard varieties. For that reason, the importance of non-standard evidence has been pointed out throughout the paper. Murelli's observations were then checked against the known rules on the use of InvRels in both standard and non-standard Croatian. It has been shown that in standard Croatian, tendencies displayed in other European languages hold very well. It was, however,



shown that standard Croatian being in line with these tendencies has more to do with language-internal rules, specifically, the clitic status of some resumptives. As for non-standard Croatian, several rules on use of InvRels different from standard Croatian were presented. The two most important ones were the accessibility of lower positions relativized, and the possibility of the omission of resumptive pronouns with DO-relativization. Both of these observed differences called for a quantitative analysis using corpus data, which was in turn provided in the second part of the paper.

For the purposes of the analysis of non-standard Croatian, only the InvRels *šta* and *kaj* were chosen. The presented corpus study has given us some important insights into the functioning of the RCs with these two InvRels. We have seen that the predominant positions relativized are SUB and DO, whereas IO and OBL are only rarely found. The omission of resumptive pronouns with DO-INAN, another important feature, has proved to be predominant in our sample. Furthermore, our data suggests that the omission of resumptives with DO-ANIM is possible but improbable. The parallel examination of the two non-standard relativizers highlighted some differences in both of the examined topics, the InvRel *kaj* being somewhat more frequent with DO-relativization and more likely to omit resumptive pronouns with DO-relativization.

The second part of the paper has once again highlighted the importance of in-depth corpus studies when relativization strategies are concerned. Our study has shown that the mere fact that an example of a certain phenomenon is attested does not mean that it is widely used. IO and OBL-relativization in non-standard Croatian was a case in point. What is more, the study has shown that the rules governing some aspects of RC formation, e.g. use of resumptive pronouns, can be established only when investigated quantitatively.

Finally, it should be noted that the analyzed sample, while being representative for non-standard, cannot completely replace a sample of spoken language. That being said, the next step in the research of relative clauses with the invariant relativizer in Croatian will have to deal with an analysis of data taken from everyday conversation.

References

- Ammon, Ulrich. 2004. Standard variety. In Ammon, Ulrich, & Dittmar, Norbert; & Matheier, Klaus J. (eds.), *Sociolinguistics: An international handbook of the science of language and society* (Vol. 1), 273–283. Berlin: Mouton de Gruyter,



- Auwers, Johan van der, & Kučanda, Dubravko. 1985. Pronoun or conjunction – the Serbo-Croatian invariant relativizer *što*. *Linguistics* 23: 917–962.
- Biber, Douglas, & Johansson, Stig, & Leech, Geoffrey, & Conrad, Susan, & Finegan, Edward. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- Browne, Wayles. 1986. *Relative clauses in Serbo-Croatian in comparison with English*. Zagreb: Institute of Linguistics, Faculty of Philosophy, University of Zagreb.
- Cheshire, Jenny, & Stein, Dieter (eds.). 1997. *Taming the vernacular: From dialect to written standard language*. London: Longman.
- Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd ed. Chicago: The University of Chicago Press.
- Comrie, Bernard, & Kuteva, Tania. 2013a. Relativization on subjects. In Dryer, Matthew S., & Haspelmath, Martin (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/122>, accessed on Feb 29, 2016).
- Comrie, Bernard, & Kuteva, Tania. 2013b. Relativization on obliques. In Dryer, Matthew S., & Haspelmath, Martin (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/123>, accessed on Feb 29, 2016.)
- Cristofaro, Sonia, & Giacalone Ramat, Anna. 2007. Relativization strategies in the languages of Europe. In Ramat, Paolo, & Roma, Elisa (eds.), *Europe and the Mediterranean as linguistic areas: Convergences from a historical and typological perspective*, 63–93. Amsterdam: John Benjamins.
- Fiorentino, Giuliana. 2007. European relative clauses and the uniqueness of the relative pronoun type. *Italian Journal of Linguistics* 19(2). 263–291.
- Fried, Mirjam. 2010. Accusative resumptive pronoun in the Czech relative clauses with absolutive relativizer *co*. *Korpus, Gramatika, Axiologie* 1(1). 16–29.
- Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. Haspelmath, Martin, & König, Ekkehard, & Oesterreicher, Wulf, & Raible, Wolfgang (eds.), *Language typology and language universals*, 1492–1510. Berlin: De Gruyter.
- Keenan, Edward, & Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.
- Kordić, Snježana. 1995. *Relativna rečenica*. Zagreb: Hrvatsko filološko društvo; Matica hrvatska.
- Langston, Keith, & Peti-Stantić, Anita. 2014. *Language planning and national identity in Croatia*. New York: Palgrave Macmillan.
- Ljubešić, Nikola, & Klubička, Filip. 2014. {bs,hr,sr}WaC – Web corpora of Bosnian, Croatian and Serbian. *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*. Gothenburg, Sweden. (Available online at <http://nlp.ffzg.hr/data/publications/>)

- nljubesi/ljubescic14-bs.pdf, accessed on Feb 29, 2016).
- Marle, Jaap van. 1997. Dialect versus standard: nature versus culture. In Cheshire, Jenny, & Stein, Dieter (eds.), 13–34.
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin: De Gruyter Mouton.
- Polančec, Jurica, & Gnjatović, Tena. 2014. The indeclinable relativizer *što* – an analysis of examples from contemporary spoken Croatian. In Peti-Stantić, Anita, & Stanojević, Mateusz-Milan (eds.), *Language as information. Proceedings from the CALS conference 2012*, 193–213. Frankfurt am Main: Peter Lang.
- Pranjković, Ivo. 1986. Koji i što. *Jezik* 34. 10–17.
- Pranjković, Ivo. 2010. Zamjenica, prilog, čestica i veznik ‘što.’ In Badurina, Lada, & Bačić-Karković, Danijela (eds.), *Riječki filološki dani: zbornik radova s Međunarodnoga znanstvenog skupa Riječki filološki dani, održanoga u Rijeci od 6. do 8. studenoga 2008*, 401–408. Rijeka: Filozofski fakultet.
- Quirk, Randolph, & Greenbaum, Sidney, & Leech, Geoffrey, & Svartvik, Jan. 1985. *A comprehensive grammar of the English language*. New York: Longman.
- Silić, Josip, & Pranjković, Ivo. 2005. *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. Zagreb: Školska knjiga.
- Šojat, Antun, & Barac-Grum, Vida, & Kalinski, Ivan, & Lončarić, Mijo, & Zečević, Vesna. 1998. *Zagrebački kaj: govor grada i prigradskih naselja*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Sonnenhauser, Barbara. 2013. Relative clauses in Slovene: diachronic puzzles, synchronic patterns. *Vienna Slavic Yearbook* 1. 150–187.
- Stein, Dieter. 1997. Syntax and varieties. In Cheshire, Jenny, & Stein, Dieter (eds.), 35–50.
- Weinert, Regina. 2004. Relative clauses in spoken English and German – their structure and function. *Linguistische Berichte* 197. 3–51.

Authors' address:

Jurica Polančec
Faculty of Humanities and Social Sciences
University of Zagreb
Ivana Lučića 3
10000 Zagreb
Croatia
E-mail: jpolance@ffzg.hr

Ana Mihaljević
Dragutina Golika 38
10000 Zagreb



Croatia

E-mail: amihaljevi@gmail.com

O NESKLONJIVIM RELATIVIZATORIMA U NESTANDARDNOM HRVATSKOM IZ AREALNE PERSPEKTIVE

U radu se razmatra uporaba nesklonjivih relativizatora u hrvatskome jeziku te je se uspoređuje s uporabom nesklonjivih relativizatora u ostalim europskim jezicima. U radu se naglasak stavlja na uporabu nesklonjivih relativizatora u nestandardnome hrvatskom. Rad je podijeljen u dva dijela. U prvome se dijelu predstavljaju već istražena pravila o uporabi nesklonjivih relativizatora u standardnome i nestandardnome hrvatskom. Pokazuje se da su nesklonjivi relativizatori običniji u nestandardnim idiomima te da su podložni manjem broju sintaktičkih ograničenja nego što je to slučaj u standardnome jeziku. Ukratko se razmatra i sociolingvistička situacija. U drugome se dijelu rada s pomoću mrežnog korpusa *hrWac 2.0* analiziraju dvije dosad neistražene pojave vezane za uporabu nesklonjivih relativizatora u nestandardnome hrvatskom: udio pojedinih položaja antecedenta u relativnoj rečenici te ispuštanje resumptivne zamjenice kada je položaj antecedenta u relativnoj rečenici izravni objekt. Zbog prirode korpusa i nekih drugih razloga istraživani su samo relativizatori *šta* i *kaj*. Pokazuje se da je najčešći položaj antecedenta subjekt te u manjoj mjeri izravni objekt, a ostali su položaji vrlo rijetki. Nadalje, pokazuje se da se resumptivna zamjenica redovito ispušta kada je antecedent što neživo, a ako je antecedent što živo, gotovo nikada.

Ključne riječi: relativna rečenica; nestandardni varijeteti; hrvatski jezik; nesklonjivi relativizator; korpusna analiza.