Competition between 'Who' and 'Which' in Slavic Light-Headed Relative Clauses*

Конкуренция местоимений 'кто' и 'который' в славянских относительных предложениях без существительного в вершине

Maria A. Kholodilova

Institute for Linguistic Studies of the
Russian Academy of Sciences;
National Research University
Higher School of Economics
St. Petersburg, Russia

Мария Александровна Холодилова

Институт лингвистических
исследований РАН; Национальный
исследовательский университет
"Высшая школа экономики"
С.-Петербург, Россия

Abstract

The relativization systems of most Slavic languages include relative pronouns that can be conventionally labelled as 'who' and 'which' and differ in a number of logically independent parameters (etymology, animacy, grammaticality of attributive contexts, and morphological distinction for number and gender). Prior research has shown that the choice between 'who' and 'which' in Slavic languages is largely dependent on the head type. Some of the languages allow the 'who' pronouns to be used with pronominal heads, but not with nouns in the head, while in others, the pronominal heads in the plural are also ungrammatical with the pronoun 'who.'

The present study aims to complement the available qualitative data on the distribution of the relativizers with quantitative data and to propose a unified account for all the observed tendencies. A corpus-based study was conducted in order to establish language-internal statistical tendencies comparable to the known grammaticality restrictions. The results show much agreement between the qualitative and quantitative tendencies. Thus, the head 'those,' unlike the head 'that,' is incompatible with the relativizer 'who' in Slovak, Polish, Upper Sorbian, and Lower Sorbian languages, while the same tendency is quantitative in Czech, Slovene, Serbo-Croatian, Ukrainian, Belarusian, and the older varieties of Russian. Corpus data suggest that there is also a stronger tendency for the relative pronoun 'who' to be avoided with the head 'those' than with the head 'all.' One more relevant parameter is the semantic type of the clause, maximizing semantics being the preferred option for 'who.' I suggest that all these and some other tendencies can be subsumed under a macroparameter of the extent to which the head is integrated into the relative clause.

Keywords

Slavic languages, relativization, relative pronouns, interrogative pronouns, light-headed relatives, free relatives

Резюме

В большинстве славянских языков система релятивизации включает относительные местоимения типов 'кто' и 'который', противопоставленные по ряду признаков (этимология, одушевленность, допустимость атрибутивных контекстов, состав парадигмы). Из существующих исследований известно, что на выбор между этими местоимениями во многом влияет тип вершины, в частности, местоимения типа 'кто' более допустимы в некоторых языках при вершинах, выраженных местоимениями ('те', 'все', 'тот' и др.), а из них — при вершинах в единственном числе. Целью настоящего исследования является дополнение известных тенденций по распределению релятивизаторов количественными данными и обобщение доступных данных по допустимости и частотности относительных местоимений в рамках единого подхода. Для получения количественных данных, сопоставимых с известными различиями по допустимости, статистически значимые количественные различия устанавливаются на материале корпусов славянских языков. Согласно проведенному исследованию, количественные и качественные тенденции по распределению местоимений 'кто' и 'который' хотя бы частично подчиняются общим закономерностям. Так, в словацком, польском и лужицких языках вершина 'те' в отличие от вершины 'тот' не допускает использования местоимения 'кто', тогда как в чешском, словенском, сербо-хорватском, украинском, белорусском и в русском языке XVIII века та же тенденция проявляется в количественных предпочтениях. Кроме того, корпусные данные позволяют установить различие между вершинами 'все' и 'те', которое не приводит к грамматическим запретам ни в одном из исследованных языков. Еще одним параметром, влияющим на распределение относительных местоимений, оказывается семантический тип относительной клаузы: местоимение 'кто' наиболее предпочтительно в максимализирующих относительных предложениях и наименее предпочтительно в нерестриктивных. Можно предположить, что эти и некоторые другие тенденции в дистрибуции местоимений 'кто' и 'который' определяются одним макропараметром — степенью вложенности вершины в относительную клаузу.

1. Introduction

Slavic languages exhibit considerable variation in the use of relativizers, which manifests itself both language-internally and cross-linguistically, see among others [Bauer 1967; Бауэр 1967; Křížková 1970; Gołąb, Friedman 1972; Murelli 2011]. A particular problem within the variation system concerns the competition of the relative pronouns that can be labelled as 'who' and 'which.' The former label is meant to cover the relativizers which directly continue interrogative pronouns of the protolanguage with the meaning 'who,' refer to people, do not decline for number and gender, and cannot be used attributively. The latter label refers to relativizers of a different origin, applicable to both animate and inanimate objects, which decline for number and gender, and can, under some conditions, be used attributively. Relative pronouns of these classes are present in most Slavic languages [Gołąb, Friedman 1972].

Language-internally, the distinction between 'who' and 'which' gives rise to a competition between syntactic synonyms, as illustrated in (1) by two Serbian translations from the same Russian original. The translation in (1a) contains the relative pronoun *koji* 'which,' while a similar meaning in (1b) is rendered by *ko* 'who.'

(1)   Mikhail Bulgakov. The Master and Margarita (Parasol)

Serbian[1]

a.   I   upravo   tu   Ivan Nikolajevič   konačno   izgubi   onoga
     and   exactly   here   Ivan Nikolaevich   finally   lost   him

     koji   *mu*   *je bio*   *toliko*   *potreban.*
     which   for.him   was   so   necessary

b.   I   eto   tu   Ivan Nikolajevič   konačno   izgubi   onoga
     and   this   here   Ivan Nikolaevich   finally   lost   him

     ko   *mu*   *je bio*   *toliko*   *potreban.*
     who   for.him   was   so   necessary

     'And it was here that Ivan Nikolaevich definitively lost him whom he needed so much.'

---

[1]   All originally Cyrillic examples in the paper are transliterated. Examples are given in professional translations extracted from parallel corpora wherever possible.

Cross-linguistically, the variation results in differences between Slavic languages with respect to the grammaticality and frequency of relative pronouns in different contexts. A preliminary idea of the variation is given by (2), where Slavic translations of a single passage contain either 'who' (2a) or 'which' (2b). The original Russian extract is given under (2a) and contains the pronoun 'who.'

(2) Mikhail Bulgakov. The Master and Margarita (RNC)

a. 'who'

te,	**kto**	*videl ego*	*vpervye. . .*	Russian
tyja,	**xto**	*bačyŭ jaho*	*ŭperšynju*	Belarusian
ty,	**kdo**	*ho viděli*	*poprvé*	Czech
those	who	saw him	first	

b. 'which'

ti,	**jaki**	*bačyly joho*	*vperše. . .*	Ukrainian
ci,	**którzy**	*widzieli go*	*po raz pierwszy*	Polish
oni	**koji**	*su ga vidjeli*	*prvi put*	Croatian
those	which	saw him	first	

'Those who saw him for the first time. . .'

The distribution of Slavic 'who' and 'which' has been surveyed in the literature, most notably by Křížková [1970] for all the modern standard Slavic languages, see also [Browne 1986; Kordić 1994; 1999; Mitrović 2012; Павловић 2012] for Serbo-Croatian, [Zubatý 1918; Svoboda 1967; Karlík 1988] for Czech, [Darovec 1985: 109–111; Podhajecká 2010; eadem 2012] for Slovak, [Polański 1967: 78–79; Fasske, Michalk 1981: 615–617, 626–627] for Upper Sorbian, [Janaš 1976: 184–185] for Lower Sorbian, and [Зализняк, Падучева 1975; Spencer 1993: 38–41; Никунласси 2008] for Russian. However, most of these studies, with the notable exception of [Kordić 1994; idem 1999: 196–197], are based only on qualitative data. To the best of my knowledge, no prior study has attempted to consider this whole subsystem of variation between relativizers in view of both qualitative and quantitative data. At the same time, such data could shed some light on the consistency of relative clause patterning across Slavic languages.

The primary objective of the present study is therefore to bring together the newly acquired quantitative data and the available qualitative data. The competition between 'who' and 'which' is most consistently observed in light-headed relatives, i.e., relative constructions without a noun in the head,

whether overt or elided.[2] I will therefore primarily confine the discussion to this class of relative constructions, even though some remarks on a more general picture will also be given.

The study needs a consistent definition of the notions 'who' and 'which,' which is introduced in **part 2**. I go on to sketch the qualitative data on the distribution (**part 3**) and present the quantitative data (**part 4**). A discussion follows in **part 5** to specify the observed tendencies. In **part 6**, I summarize the tendencies and propose a unifying macro-parameter of variation, namely the extent to which the head is integrated into the relative clause.

2. The definition of 'who' and 'which' pronouns

As stated in the introduction, I base the distinction between 'who' and 'which' on several properties. The preliminary list of properties and sources of information follows below:

1) whether the relativizer directly continues an animate interrogative pronoun ('who') of a proto-language (Proto-Slavic *kъto in the case of Slavic languages) [Vaillant 1958; Bauer 1967; Трубачев 1987: 125];

2) synchronic semantic restrictions on its use with non-human referents [Křížková 1970];

3) the presence of number and gender distinctions in its paradigm (individual grammatical descriptions);

4) the grammaticality of attributive uses [Křížková 1970].

It is important to note that the terms 'who' and 'which' are conventional labels for comparative concepts in the sense of [Haspelmath 2010] and need not reflect the properties of the English pronouns *who* and *which*.[3]

Two common Slavic patterns with respect to the properties listed above are presented in Table 1. I preliminarily define 'who' and 'which' pronouns as those which adhere to these patterns.[4]

---

[2] The term "light-headed relatives" for the relative clauses without a noun in the head was coined by Citko [2000; 2004]. Another common term for this group of relative constructions is "false free relatives" [De Vries 2002].

[3] The English *who* does happen to fall into the 'who' group, whereas *which* does not quite correspond to the set of features of the 'which' group because it does not have distinct plural forms.

[4] An anonymous reviewer suggests Slavic possessive relative pronouns ('whose') should be regarded as possessive forms of the 'who' pronouns. My data suggest that these pronouns consistently show different syntactic properties and different patterns of use, see among others [Беличова 1988; Kholodilova 2013]. They are therefore not included in this survey.

Table 1. The differences between the typical Slavic 'who' and 'which'

	stems from the interrogative 'who'	human reference only	does not have number or gender distinctions	cannot be used attributively
pattern 1 (≈ 'who')	+	+	+	+
pattern 2 (≈ 'which')	−	−	−	−

Although the proposed definition is new, the distinction itself is commonly drawn in papers specifically addressing relative clauses [Křížková 1970; Зализняк, Падучева 1975] as well as general grammatical descriptions of Slavic languages and is often construed as that between adjectival and substantive relativizers. I follow [Bauer 1967: 301] and [Křížková 1970: 13–17] among others in classifying Slavic relative pronouns in the manner reflected in Tables 2 and 3, and I attempt only to articulate the underlying intuition more precisely.

Two additional remarks are due.

First, a straightforward classification according to etymological origin is problematic for Upper Sorbian *štóž* 'who,' Bulgarian *kojto* 'which,' and Macedonian *koj* 'which' and *kojšto* 'which.' All these pronouns retain the etymological root which directly continues **kъto* only in (some of) the forms other than the nominative [Vaillant 1958: 411, 417]. This complex etymology is reflected in Tables 2 and 3 as "+/−" for Upper Sorbian *štóž* 'who,' which does not have any restrictions specific to the use of the non-nominative forms, and "−/+" for Bulgarian and Macedonian pronouns, which only use this stem, unlike the nominative forms, in animate contexts. In other words, if these Bulgarian and Macedonian pronouns are viewed as single items, it is natural to assume that the non-animate accusative forms are either derived from the *koj* stem or lacking altogether, and in either case the etymologically animate stem is only used in a subclass of oblique forms. In all cases of an intermediate status, I consider this property non-decisive for the classification of the pronouns and rely on other features.[5]

Second, among the relativizers with the properties of 'which' pronouns, I consider in this study only those with high frequency.

Tables 2 and 3 summarize the data on the definitional properties of Slavic pronouns. I consider here only the modern standard varieties of the languages, while some deviations from these descriptions are reported for dialects and older varieties.

---

[5] Note that this criterion is etymological rather than synchronical. Bulgarian, Macedonian, and Serbo-Croatian pronouns classified here as 'which' can be used in the contexts of the interrogative 'who' in the modern language, but do not directly continue the Proto-Slavic **kъto*.

Table 2. Slavic 'who' pronouns

			only human	stems from the interrogative 'who'		no number or gender distinctions	cannot be used attributively
				= 'who'	= 'who' + smth.		
East Slavic	Russian	*kto*	+	+		+	+
	Ukrainian	*xto*	+	+		+	+
	Belarusian	*xto*	+	+		+	+
West Slavic	Polish	*kto*	+	+		+	+
	Czech	*kdo*	+	+		+	+
	Slovak	*kto*	+	+		+	+
	Upper Sorbian	*štóž*	+ [6]		+/−	?	+
	Lower Sorbian	*chtož*	+		+	? [7]	+
South Slavic	Slovene	*kdor*	+		+	+	+
	Serbo-Croatian	*(t)ko*	+	+		+	+

Table 3. Slavic 'which' pronouns

			only human	stems from the interrogative 'who'	no number or gender distinctions	cannot be used attributively
East Slavic	Russian	*kotoryj*	−	−	−	−
	Ukrainian	*kotryj*	−	−	−	−
		jakyj	−	−	−	−
	Belarusian	*katory*	−	−	−	−
		jaki	−	−	−	−
West Slavic	Polish	*który*	−	−	−	−
	Czech	*který*	−	−	−	−
	Slovak	*ktorý*	−	−	−	−
	Upper Sorbian	*kotryž*	−	−	−	−
	Lower Sorbian	*kótaryž*	−	−	−	−
South Slavic	Slovene	*kateri*	−	−	−	−
	Serbo-Croatian	*koji*	−	−	−	−
	Macedonian	*koj*	−	−/+	−	−
		kojšto	−	−/+	−	−
	Bulgarian	*kojto*	−	−/+	−	−

---

[6] According to [Bartels, Spiess 2012], the relative 'who' can refer to inanimates in older literary Lower Sorbian and non-standard Upper Sorbian, unlike the present-day standard languages.

[7] Mucke [1891: 428] provides a paradigm of the interrogative *chto*, which includes instrumental plural and dual forms; however, these forms are not mentioned by Janaš [1976: 179].

Macedonian *koj* and *kojšto*, as well as Bulgarian *kojto*, do not have a counterpart belonging to the 'who' group and are therefore not considered below in any detail.

As is evident from the data in Tables 2 and 3, the four properties strongly tend to pattern together in Slavic languages. It is therefore impossible to tear them apart, building on the Slavic data only. Although a comparative concept based on a single property would be more desirable, choosing any one of the properties listed above in this study would be largely arbitrary and could yield incorrect conclusions.

3. Qualitative data on the distribution of Slavic 'who' and 'which'

To the best of my knowledge, the most systematic overview of the distribution in question is provided by Křížková [1970], who describes the use of 'who' and 'which' in questions (3), correlatives (4), and postnominal relative constructions headed by the pronouns 'someone,' 'nobody,' 'each,' etc. (5), 'that' (6), 'all' or 'those' (7), and nouns (8). The paper also contains some data on constructions with 'first,' 'last,' 'that' in the feminine, and personal pronouns in the head, as well as cleft constructions; the evidence of this data, however, is more sketchy, and I leave it out in the following discussion.

(3) question, Carlo Collodi. Pinocchio (RNC)

Kto	*vas*	*oskorbljaet?*			Russian
Kto	*wam*	*ubliża?*			Polish
Tko	*vas*	*vrijeđa?*			Croatian
who	you	insults			

'Who is insulting you?'

(4) correlative, Ivo Andrić. The Bridge on the Drina (Parasol)

Xto	*ne bačyc'*	*zaraz,*	*toj*	*ne pabačyc' nikoli.*	Belarusian
Kto	*nevidí*	*teraz,*	*ten*	*neuvidí nikdy.*	Slovak
Ko	*ne vidi*	*sada,*	*taj*	*neće nikad vidjeti.*	Serbian
who	doesn't see	now	that	will never see	

'The one who does not see now will never see.'

(5) 'each' in the head, Ivo Andrić. The Bridge on the Drina (Parasol)

. . . *Kožnomu,*	***xto***	*joho*	*xotiv*	*sluxaty*	. . .	Ukrainian
. . . *každému,*	***kdo***	*ho*	*byl ochoten*	*poslouchat* . . .		Czech
. . . *vsakomur,*	***kdor***	*ga*	*je hotel*	*poslušati*	. . .	Slovenian
each.one	who	him	wanted	to listen		

'. . . to everyone who wanted to listen to him.'

(6) 'that' in the head, Nikolai Ostrovsky. How the Steel was Tempered (Parasol)

a. . . . *zvjarnuŭsja ën da taho,* **kaho** *zvali Moc'kam*　　　Belarusian

　　. . . *obratise on onome* **koga** *su zvali Moćka*　　　Croatian

　　turned he to the one whom called Motka

b. . . . *přiwobroći so k tomu,* **kotrehož** *běchu Motku wołali*　　Upper Sorbian

　　. . . turned to the one which called Motka

'He addresses the one who was called Motka.'

(7) 'those' in the head, Umberto Eco. The Name of the Rose (Parasol)

a. *I te,* **kto** *ubival obezumevšix kajal'ščikov. . .*　　Russian

　　and those who killed crazed penitents

b. *A ci,* **którzy** *zabijali oszalałych pokutników. . .*　　Polish

　　A oni **koji** *su ubijali pomahnitale pokajnike. . .*　　Serbian

　　and those which killed crazed penitents

'And they who killed the crazed penitents. . .'

(8) noun in the head, Dan Brown. The Da Vinci Code (RNC)

　　. . . *vrag,* *o* **kotorom** *ee predupreždali*　　Russian

　　. . . *nepriateľ,* *pred* **ktorým** *ju vystríhali*　　Slovak

　　. . . *sovražnik,* *pred* **katerim** *so jo posvarili*　　Slovene

　　enemy about which they warned her

'. . . the enemy they had warned her about'

Table 4 summarizes Křížková's [1970] data. Some of the data points are corrected and highlighted in bold.[8]

First, *pace* Křížková, Upper Sorbian does allow 'who' to be used with indefinites [Fasske, Michalk 1981: 616], cf. a corpus example:

(9) Upper Sorbian, Nikolai Ostrovsky. How the Steel Was Tempered (Parasol)

Ale bydlenje běše prózdne, a **nichtó** *tu njeběše,* **kohož** *by so woprašał*

but flat was empty and nobody there wasn't whom would ask

'But the flat was empty, and there was nobody to ask.'

Second, corpus data show that Lower Sorbian allows 'which' with the head 'that':

---

[8]　Křížková [1970] sometimes does not state explicitly that one of the pronouns is ungrammatical in a given context, but she gives lists of languages using different strategies which could be expected to be exhaustive.

Table 4. The use of 'who' and 'which' in Slavic languages,
according to the data in Křížková [1970] with minor corrections

	questions	correlatives	'someone,' 'nobody,' 'each'	'that'	'all,' 'those'	nouns in the head
Russian	*who*	*who/which*	*who*	*who/which*	*who/which*	*which/[?]who*
Ukrainian	*who*	*who*	*who*	*who/which*	*who/which*	*which/[?]who*
Belarusian	*who*	*who*	*who*	*who/which*	*who/which*	*which/[?]who*
Czech	*who*	*who*	*who*	*who/which*	*who/which*	*which*[9]
Slovene	*who*	*who*	*which*	*who/which*	*who/which*	*which*
Polish	*who*	*who*	*who*	*who/which*	*which*	*which*
Lower Sorbian	*who*	*who*	*who/which*	*who/which*	*which*	*which*
Slovak	*who*	*who*	*who/which*	*who/which*	*which*	*which*
Upper Sorbian[10]	*who*	*who/which*	*who/which*	*who/which*	*which*	*which*
Serbo-Croatian[11]	*who/which*	*who/which*	*who/which*	*who/which*	*who/which*	*which*

(10) Lower Sorbian, James 2:13 (DOTKO)[12]

*Njeſmilny ßud ſmějo ten, **kótaryž** ſmilnoſcź ńejo zynił*
Merciless judgment will.have that which mercy didn't do

'For judgment is without mercy to the one who has shown no mercy.'

Third, the data on the heads 'somone,' 'nobody,' and 'each' in Lower Sorbian are missing from [Křížková 1970]. Both 'who' and 'which' pronouns are attested in such contexts in DOTKO. Cf. also the following example:

(11) Lower Sorbian, Antoine de Saint-Exupéry. The Little Prince (ASPAC)

*... mimo někogo, z **kótarymž** by mógał se derje wulicowaś*
... without someone with which would be able well talk

'without anyone that I could really talk to'

---

[9] Czech examples of 'who' pronouns with a noun in the head can be found, but they are extremely infrequent.

[10] 'Who' is attested more widely in older literary Lower Sorbian and non-standard Upper Sorbian [Bartels, Spiess 2012: 227].

[11] Browne's [1986: 34] judgments on Serbo-Croatian are in many respects different.

[12] Lower Sorbian writing in DOTKO texts deviates from the modern orthography; see a description of the correspondences at http://dolnoserbski.de/korpus/psawopisne_warianty.

Fourth, Slovene examples with 'which' following the head 'all' are also attested:[13]

(12) Slovene (Fida Plus)

*Vsekakor pa vsi, **kateri** so sodelovali pri popravilu, zaslužijo pohvalo!*
anyway but all which participated in repairs deserve praise
'Anyway, everyone who took part in the repair deserves praise!'

Most of the data in the table is consistent with the hierarchy in (13), the positions to the left being increasingly preferred for the 'who' pronouns.

(13) questions > correlatives > 'someone,' 'nobody,' 'each'; 'that' > 'all,' 'those' > heads with nouns

The only exception to this rule is represented by Russian correlatives, which unexpectedly allow the pronoun 'which.' Three comments are due here. First, standard Russian does not allow 'which' in any of the relative constructions without a noun in the head, i.e., standard Russian data clearly fits into the hierarchy in (13). Second, 'which' in non-standard Russian also occurs in the "intermediate" group, namely, with the head 'each,' in a clearly non-elliptical context, and these examples are relatively acceptable, at least when the head is feminine, as in (14).[14]

(14) Russian, T. Putilova ("Komsomol′skaya Pravda," 2002)

*Pravda, po ego slovam, on ženilsja na každoj, **kotoruju** celoval.*
truth according.to his words he married on each.F which kissed
'However, according to him, he married every girl he kissed.'

Third, even though correlatives with 'which' are slightly more acceptable and frequent than this intermediate step, it is only true for the correlatives with the relative pronoun in the plural. However, comparable conditions are impossible for the group 'someone,' 'nobody,' and 'each,' and the data discussed in 5.2 show that plural number is an independent parameter which makes 'which' pronouns more acceptable.

It is therefore possible to suggest that the hierarchy in (13) only holds if other things are equal, i.e., if the features unspecified in the group descriptions coincide.

---

[13] Slovak and Polish examples of this kind can also be found, although they are extremely rare.

[14] The examples with 'nobody' and 'someone' are, however, ungrammatical, which suggests that this group is not quite homogeneous. I will not go into further details in this paper.

The greatest variation in the table is observed in the 'that' and 'all,' as well as 'those' columns, which is why further discussion is mainly confined to this class of light-headed relatives.

4. Quantitative data on the distribution of Slavic 'who' and 'which'

The basic idea behind this study is to propose a unified account for qualitative and quantitative data on the distribution of 'who' and 'which.' The desired quantitative data point must therefore be directly comparable to the qualitative data in Table 4. It is possible to establish such comparative data by looking at information on the presence of a quantitative tendency established language-internally, which can be compared to the qualitative tendencies summarized in (13).

A potentially attractive option for a quantitative cross-linguistic study is to increase comparability of the data by using parallel corpora, such as ASPAC, InterCorp, the parallel subcorpus of the RNC, or Parasol, as advocated for Slavic languages by von Waldenfels [2014]. However, this methodology seems less suited for this study. Using parallel texts does not seem to make the data of basic interest for this study any more reliable, specifically, the data on language-internal quantitative tendencies (or, to put it in other words, the differences between columns in Table 4). It could be more useful if the intended results included not only a comparison of contexts, but also a quantitative comparison of languages (i.e., differences between rows in Table 4). However, this aim does not seem to be fully achievable. Only one text in all the Slavic varieties considered in this study can be currently found in the corpora mentioned above. This is largely due to the fact that very few texts are available in Lower Sorbian. Moreover, even if Lower Sorbian is excluded from consideration, we still end up with only four texts. The constructions in question are relatively infrequent; therefore the small amount of data could simply be non-sufficient. To give one example, the four texts mentioned above with multiple translations contain only a total of nine examples of Russian relative clauses with the head 'all' immediately preceding relative pronouns.[15] Moreover, the longest of these texts is of Russian origin, which would also be a major problem for this study. For instance, it is plausible that the sentences originally in Russian would be more likely to be calqued in languages which are closer to it structurally and lexically. Finally, if only a few translated texts are taken into account, it is likely that the differences "between rows of Table 4" will reflect the individual translators' preferences rather than the intrinsic properties of the languages.

---

[15] I only consider the first Russian translation when more than one variant is available.

To sum up, consistent use of translated texts would not provide any advantages for the core of this study and could in some respects be harmful. I therefore do not intend to collect directly comparable quantitative data on the absolute frequency of relativizers in different languages, but only the data on the tendencies of their use. This implies, among other things, that not only the corpora and texts, but also the principles of the search, such as the search query and sampling, need not be exactly the same across the languages, as long as there is no independent reason to expect that these features can reverse the correlations.

Table 5 presents the data obtained in this study.

The first column corresponds to the surveyed languages. I do not consider modern Russian data, because *kotoryj* 'which' in modern standard Russian is used almost exclusively with nouns in the head, see, e.g., [Spencer 1993]. Surveying non-standard data poses additional problems related to the distinction between light-headed relatives and relatives with elided heads, which are extremely frequent in non-standard texts. However, to fill this gap, the data on Russian of the 18th century is included.[16]

The second column lists the pronouns in the head that were included in the search queries.[17]

The columns 'who' and 'which' contain the raw frequencies of the corresponding constructions. In all cases, the relative pronouns were included in the queries in the nominative. All the search results were inspected manually in order to exclude the relative clauses with elided nouns in the head. The next column conveys the ratio the 'who' constructions constitute among the relatives of both types.

The last column lists the corpora used in the substudies.

---

[16] The legitimacy of including an earlier language variety in the sample, which was questioned by anonymous reviewers, derives from the following assumption. The aim of this study is to propose a generalization that would hold true for as many language varieties as possible. I therefore build a sample of language varieties which are, first, suitable for this study and, second, as distantly related as possible. In most cases, this sample happens to coincide with the list of modern standard Slavic languages, because they are distantly related and have large corpora. In Russian, however, the modern language does not satisfy the first condition, and therefore another variety must be considered. The choice of the century is largely random. A supplementary micro-diachronic study shows a gradual decline in the frequency of the 'which' pronoun in this construction over the last three centuries, and the end of the 18th century is a cut-off point as good as any other.

[17] The preferable option was to search only for the heads in the nominative; however, this would not be sufficient for Slovene, in which both constructions under study are extremely infrequent, whereas the preferred option is the indeclinable relativizer. Moreover, the nominative forms were not considered in the Slovene sample of demonstrative pronouns because they do not differentiate between singular and plural.

Table 5. Quantitative data on the use of 'who' in different contexts

	head	'who'	'which'	ratio of 'who'	Corpora
Belarusian	*toj* 'that'	180	8	1.0	RNC, parallel subcorpus
	use 'all'	43	1	1.0	
	tyja 'those'	105	12	0.9	
Ukrainian	*toj* 'that'	64	1	1.0	
	usi / vsi 'all'	16	0	1.0	
	ti 'those'	32	11	0.7	
Slovene	*tisti* 'that'	10	0	1.0	FidaPLUS
	vsi 'all'	4	4	0.5	
	tisti 'those'	0	9	0.0	
Czech	*ten* 'that'	91	9	0.9	InterCorp
	všichni 'all'	78	22	0.8	
	ti 'those'	54	46	0.5	
Russian of the 18th c.	*tot* 'that'	206	83	0.7	RNC, main subcorpus
	vse 'all'	3	7	0.3	
	te 'those'	6	161	0.0	
Serbo-Croatian[18]	*onaj* 'that'	16	17	0.5	InterCorp
	svi 'all'	0	10	0.0	
	oni 'those'	0	94	0.0	
Slovak	*ten* 'that'	61	14	0.8	
	všetci 'all'	0	11	0.0	
	tí 'those'	0	33	0.0	
Polish	*ten* 'that'	53	47	0.5	
	wszyscy 'all'	0	100	0.0	
	ci 'those'	0	100	0.0	
Lower Sorbian	*ten* 'that'	4	20	0.2	DOTKO
	wšykne 'all'	0	9	0.0	
	te 'those'	0	71	0.0	
Upper Sorbian	*tón* 'that'	0	100	0.0	HOTKO
	wšitcy 'all'	0	100	0.0	
	ći 'those'	0	100	0.0	

---

[18] Serbian and Croatian were surveyed separately; however, the results did not show any difference. The results in the table are taken from the Croatian subcorpus.

Table 6 is a reminder of the differences between 'that,' 'those,' and 'all' observed in the qualitative data. The vertical lines mark the cases where 'who' is acceptable in one of the contexts, but not the other. The relativizer 'which' is attested in all these languages with all three head types.

Table 6. Grammaticality of 'who' in different contexts, according to the data in Křížková [1970] with minor corrections

	'those who,' 'all who'	'that who'
Belarusian	+	+
Ukrainian	+	+
Slovene	+	+
Czech	+	+
Russian	+	+
Serbo-Croatian	+	+
Slovak	−	+
Polish	−	+
Lower Sorbian	−	+
Upper Sorbian	−	+

Table 7 summarizes the results of the quantitative study. The numbers in the cells correspond to the ratio of 'who' in these contexts. The dashed vertical lines designate statistically significant differences between two contexts with $p < .05$.[19] The normal lines are copied from Table 6. The last column reflects statistically significant differences between the leftmost ('that') and the rightmost ('those') columns.

Table 7. Quantitative and qualitative data on the use of 'who' in different contexts. Summary (ratio of 'who')

	'that who'	'all who'	'those who'	statistically significant difference between 'that' and 'those'
Belarusian	1.0	1.0	0.9	+
Ukrainian	1.0	1.0	0.7	+
Slovene	1.0	0.5	0.0	+
Czech	0.9	0.8	0.5	+
Russian of the 18th c.	0.7	0.3	0.0	+
Serbo-Croatian	0.5	0.0	0.0	+
Slovak	0.8	0.0	0.0	+ (also qualitative)
Polish	0.5	0.0	0.0	+ (also qualitative)
Lower Sorbian	0.2	0.0	0.0	+ (also qualitative)
Upper Sorbian	0.0	0.0	0.0	− (qualitative)

---

[19] Two-sided Fisher's exact test was applied if at least one of the expected values was lower than 5; χ^2 test was used in all other cases.

Two points are worth noting. First, the qualitative and quantitative data on the use of relative pronouns with 'that' and 'those' follow a common tendency for 'who' to be less preferred (either disallowed or less frequent) with the head 'those.' As shown in Table 7, this tendency is observed or was observed in **all** the standard Slavic varieties considered in this study. Second, the quantitative data suggests a further difference between 'these' and 'all,' the relativizer 'who' being more frequent in the latter case. This tendency is observed in at least four of the five Slavic languages for which the frequency of both relative pronouns with plural heads is not vanishingly small.

I assume that the two differences have a common nature, even though the second one happens not to be reflected in the grammaticality constraints of any of the languages in the study. I therefore propose the second interim summary (15) with the newly found distinction taken into account.

(15) questions > correlatives > 'someone,' 'nobody,' 'each'; 'that' > **'all'** > **'those'** > heads with nouns

5. Discussion

The hierarchy in (15) incorporates many different properties that could be responsible for the observed grammatical differences. The discussion in **parts 5.1–5.2** is intended to clarify which of them could be relevant.

5.1. Positional types of relative clauses and "lightness" of the head

At least some parts of the hierarchy in (15) can be described in terms of a diachronical cline, the positions to the left being diachronically prior.

The first step in the hierarchy (15) corresponds to a well-established grammaticalization path:

(16) interrogatives > (indefinites) > relative markers [Lehmann 1984: 384; Henderey 2012: 48–56]

One more part of the hierarchy which can be interpreted in diachronical terms is given in (17).

(17) correlatives > **light-headed relatives** > relatives with nouns in the head

This subhierarchy captures the synchronical fact that relative clauses without a noun in the head are in a sense intermediate between free relatives and relatives with full heads.

The fact that free relatives often share the relativizer with (some) relative clauses with pronouns in the head was noted on a typological basis already in [Зализняк, Падучева 1975] and [Lehmann 1984]. According to Lehmann [1984: 326], if a relative pronoun coincides with an interrogative pronoun in

noun-headed relatives, it always does so in light-headed relatives. This suggests that light-headed relatives might constitute an intermediate diachronical stage between correlatives and postnominal relative clauses.[20]

As noted in the literature, light-headed relatives have some properties in common with noun-headed relatives, such as the lack of matching effects, a similar set of semantic interpretations, and non-compatibility with *ever* [Szucsich 2003; Citko 2004]. A link between light-headed relatives and relatives with a noun in the head is also part of a proposed grammaticalization cline of relative pronouns by Heine and Kuteva [2006], see also [Mithun 2012].

5.2. Specificity, number, and semantic type

Several researchers have independently claimed that a difference relevant for the choice of relativizers in light-headed relative clauses is that between relatives with specific and non-specific reference. This idea is proposed in these or in slightly different terms (in part due to different languages) by Křížková [1970], Lehmann [1984], and Spencer [1993]. According to all these studies, 'who' is more pervasive in relative clauses with non-specific reference.

According to Křížková [1970], 'which' in Czech, Slovak, Russian, Ukrainian, and Belarusian can be used with the singular head 'that' only in definite specific contexts. As demonstrated by (17) and (18), this rule is not without exceptions, at least in Slovak and Czech, but the general tendency clearly holds.

(18) Slovak, Love and Death (Intercorp, subtitles)

Ten,	***ktorý***	*žije*	*s*	*mečom,*	*zomrie*	*mečom.*
that	which	lives	with	sword	dies	with.sword

{Political assassination doesn't work. Violence leads to violence.} 'He who lives by the sword dies by the sword.'

(19) Czech, My Name Is Bruce (Intercorp, subtitles)

Ten,	***který***	*se bude chtít utkat*		*s*	*Guan-Dim,*	—
that	which	will want to encounter		with	Guan-Di	

utká se		*také*	*se*	*samotnou*	*smrtí.*
will encounter		too	with	itself	death

'He who would confront Guan-Di,—would also confront death itself.'

---

[20] Diachronically, however, the development of Slavic relatives did not follow this scenario unidirectionally. The group of relatives with nouns in the head is attested for Slavic 'who' in a few Slavic non-modern varieties which do not allow this construction now, at least, older varieties of Serbo-Croatian [Dezső 1982: 219; Murelli 2011: 314], Middle Polish [Senderska 2013: 333], and older Lower Sorbian [Bartels, Spiess 2012: 227]. It was also more frequent in Middle Russian [Борковский 1973].

Keeping this in mind, it is natural to suppose that this tendency could account for the observed difference between 'that' and 'these.' Plural heads could just be more likely to be specific, because the singular 'that' pronouns frequently participate in constructions with generalizing semantics. To check whether specificity could be the sole relevant factor, a pilot study was conducted based on the 18th-century Russian material. The relative clauses with the head 'those' were compared to relative clauses with the head 'that,' further divided into relative constructions with specific reference (20) and relative constructions with non-specific reference (21).

(20) Russian, Nikolai Karamzin, 1793 (RNC)

*Esli   on   sam   poslal tebja — tot,   **kotorogo** strašnoe   prokljatie*
if   he   himself   sent   you   that of.which   terrifying curse

gremit   vsegda   v   moem   sluxe. . .
thunders   always   in   my   hearing

'If he sent you himself, he whose terrible curse always sounds in my ears. . .'

(21) Russian, Denis Fonvizin, 1788 (RNC)

*Basn' učit,   čto tot, **kotoryj** pervee vsex   prinimaet mody,*
fable   teaches that that which   earlier than.everyone accepts   vogue

i   tot, kotoryj deržitsja stariny, — oba   ravnye duraki.
and   that which   keeps.to   the.old   both   equal   fools

'The fable teaches us that the one who is the first to accept the new fashion and the one who keeps to the old are equally fools.'

The study was designed to check whether specificity correlates with the choice of the relative pronoun in this Russian variety and whether this difference could account for the deviant behavior of the plural heads.

In the singular, the difference was drawn between the constructions with generalized reference, classified here as non-specific, and all other cases. The difference in specificity is difficult to observe in the plural, which is why no attempt was made to classify the occurrences of the plural into these two groups.

The results are given in Table 8. The data show statistically significant differences in both pairs. The difference between specific and non-specific 'that' (χ^2, p < .05) suggests that this property was relevant for Russian of the 18th century. The difference between specific 'that' and all the occurrences of 'those' (χ^2, p < .05) suggests that even if we presume that all the examples with 'those' are specific, the number is still a different factor.

Table 8. Frequency of 'who' and 'which' in Russian of the 18th century:
The impact of number and specificity (RNC)

		'who'	'which'	ratio of 'who'
'that'	non-specific	79	21	0.79
	specific	29	71[21]	0.29
'those'		4	96	0.04

To sum up, both number and specificity influenced the choice of relative pronouns in at least one Slavic variety. As the number distinction shows up in all the Slavic varieties considered in this study, it is likely to be an independent parameter in at least some of these varieties.

I suggest that the data on number distinctions presented so far can be summarized as follows:

(22) SG > PL

(23) no (semantically regular) SG/PL distinction > SG/PL distinction

As in all the hierarchies above, the properties at the left make the choice of 'who' more favorable. The hierarchy in (22) is a proposed summary for the discussion of the difference between 'that' and 'those,' and (23) is a suggested way to account for the difference between 'all' and 'those.'[22]

Bulgarian, which is not considered in the core of this study, gives some additional data in favor of (22). As mentioned in part 2, Bulgarian does not distinguish between 'who' and 'which' in the terms of this study, and the two etymological roots constitute a single paradigm. However, according to [Ни-цолова 1986; Илиев 2012: 192], the forms derived from the former 'who' stem can only be used in **singular** masculine animate contexts, i.e., the remnants of the 'who' pronoun in Bulgarian are preferred in the singular in accordance with the hierarchy in (22).

Returning to the question of specificity, it is plausible to suggest that this distinction can be rephrased as that between maximalizing and restrictive relative clauses, as described by [Grosu, Landman 1998], maximalizing relative clauses corresponding to the non-specificity condition. As claimed in [ibid.], maximalizing relatives demand universal or definite semantics. While definiteness of examples like (20) can easily be explained by the presence of a definite pronoun in the head, the universal semantics of (21) is likely to result from a maximalizing operation. With this assumption, the generalization

---

[21] About half of the examples in this group have reference to God. This subgroup has a stronger tendency to contain the pronoun 'which.' The tendencies presented in Table 7 hold true if these examples are excluded from the survey.

[22] This generalization could also capture the colloquial Russian data discussed in footnote 14.

of the observed data might be laid out as follows: the relative pronouns of the 'who' class prefer maximalizing relative clauses over restrictive relative clauses.[23]

A further distinction between semantic types of relative clauses is drawn in East Slavic languages, where 'who' is possible in (a subclass of) restrictive relative clauses with nouns in the head and impossible in appositive relative clauses with nouns in the head.[24] As illustrated by (24)–(29), Russian *kto*, Ukrainian *xto*, and Belarusian *xto* 'who' allow only the restrictive interpretation, while *kotoryj*, *kotry*, and *jaki* 'which' are grammatical with both restrictive (24a), (26a), (28a) and appositive (24b), (26b), (28b) interpretations.

(24) Russian, M. Sergeev. The Magic Galosha (RNC)

Vzroslye, **kto** *v etot čas byl svoboden ot raboty, prinesli stul'ja. . .*
grown-ups who in this hour was free   from work   brought chairs

'The grown-ups who were free at that time brought chairs.'

a. ᴼᴷ'Some of the grown-ups were free.'

b. *'All the grown-ups were free, and the speaker knows it.'

(25) Russian, (constructed)

Vzroslye, **kotorye** *v etot čas byli svobodny ot raboty, prinesli stul'ja. . .*
grown-ups which   in this hour was free   from work   brought chairs

'The grown-ups(,) who were free at that time(,) brought chairs.'
a. ᴼᴷ'Some of the grown-ups were free.'
b. ᴼᴷ'All the grown-ups were free, and the speaker knows it.'

(26) Ukrainian, M. Frolova. The Ball in the Firmament (RNC)[25]

Zbihlys'   susidy, **xto** *buv nedaleko. . .*
came.running neighbors who was not.far
'The neighbors who were nearby came running.'
a. ᴼᴷ'Some of the neighbors were nearby.'
b. *'All the neighbors were nearby, and the speaker knows it.'

---

[23] Alternatively, both (20) and (21) can be viewed as maximalizing, universal semantics being diachronically the first step by which interrogative pronouns acquire maximalizing semantics [Belyaev, Haug]. In this case, this tendency reflects the general inclination of interrogative pronouns to be used in diachronically early contexts.

[24] An anonymous reviewer suggests that these relative clauses could be free relatives juxtaposed to the noun. The data that prove *kto*-relatives can modify nouns in the head includes the agreement of the predicate in these clauses with the head [Холодилова 2015] and the fact that the head can contain the word *tot* 'that' without the deictic meaning, which is impossible without a modifying relative clause.

[25] I would like to thank Viktor A. Stegnij for the grammaticality judgments of the Ukrainian sentences and Zhanna Duko for the judgments of the Belarusian sentences.

(27) Ukrainian (constructed)

Zbihlys'   *susidy,*   **kotri** *buly*   *nedaleko. . .*
came.running   neighbors   which   were   not.far

'The neighbors(,) who were nearby(,) came running.'

a. ^{OK}'Some of the neighbors were nearby.'

b. ^{OK}'All the neighbors were nearby, and the speaker knows it.'

(28) Belarusian, I. Chigrinov. Blood Acquittal (RNC)

Palonnyja, xto   *jašče stajaŭ*   *lja*   *scjany, pačali aziracca*   *na kryk*
captives   who   still   stood.sg near   wall   started   look.around on   cry

i   *adyxodzic' na*   *sjaredzinu dvara.*
and   step.away   on   middle   of.yard

'The captives who were still standing near the wall started to turn to the cry and move to the center of the yard.'

a. ^{OK}'Some of the captives were near the wall.'

b. *'All the captives were near the wall.'

(29) Belarusian, I. Chigrinov. Blood Acquittal (RNC)

Palonnyja, jakija jašče stajali   *lja*   *scjany, pačali aziracca*   *na kryk*
captives   which   still   stood.pl near   wall   started   look.around on   cry

i   *adyxodzic' na*   *sjaredzinu dvara.*
and   step.away   on   middle   of.yard

'The captives(,) who were still standing near the wall(,) started to turn to the cry and move to the center of the yard.'

a. ^{OK}'Some of the captives were near the wall.'

b. ^{OK}'All the captives were near the wall.'

It is therefore possible to provisionally summarize the differences between semantic types of clauses which favor the use of 'who' as follows:

(30) maximalizing > restrictive > appositive

6. Summary

The tendencies discussed in part 5 are brought together in the following schema (see Figure 1).

Most of these generalizations are based on the hierarchy (15) which is justified by the data of numerous Slavic languages. The second part of the hierarchy (34) is less reliable, because it is backed only by the data of East Slavic languages.

Figure 1

'who' ◄─────────────────────────────► 'which'

HEAD:

(31) no external head > external head without a pronoun > external head with a noun

(32) no semantically regular SG/PL distinction > SG/PL distinction

(33) SG > PL

CLAUSE:

(34) maximalizing > restrictive > appositive

The choice of relative pronoun is usually regarded as the sole argument in favor of a separate analysis of light-headed vs. noun-headed relative constructions, see, e.g., [Szucsich 2003; Citko 2004]. It follows naturally from the hierarchies related to the properties of the head that light-headed relatives do not constitute a homogeneous typologically relevant grammatical class. In other words, if this distinction is preserved, it is typologically more useful to postulate a gradual opposition between "lighter" and "heavier" heads rather than a binary opposition between light vs. noun-headed relatives.

Building on the data in (31)–(34), I would like to propose that all these differences can be derived from a single macro-parameter, namely, the extent to which the head is integrated into the relative clause. The more integrated heads are more likely to take 'who' in the relative clause.

This macro-parameter would naturally explain (34), because, according to Grosu and Landman [1998: 126], the semantic impact of the material external to the relative clause on the overall semantics of the construction declines along the following hierarchy:

(35) Simplex XPs—Appositives—Restrictives—Maximalizers—Simplex CPs

As a head gets more integrated into the relative clause it is natural to assume that it is less likely to have an independent impact on the overall semantics; see also the discussion in [Ландер 2012: 411–417].

Integration is also more natural for a head that is simply smaller segmentally and structurally (31) and has less "semantic material" (32). The difference in (33) could be regarded as that between the more and the less frequent options, which are, accordingly, less or more informative. The more informative plural form can have more impact on the semantics of the construction.

The generalization is additionally supported by four other phenomena in Russian, all of which involve some kind of interaction between the material inside the relative clause and the material outside the relative clause. These phenomena are the following:

— Agreement of the relative clause predicate with the head in Russian relative constructions [Холодилова 2015].

— Inverse attraction in non-standard Russian [Kholodilova 2015]. Under inverse attraction, the head of the relative clause is case-marked as if it belonged to the dependent clause, as in (36); see, e.g., Bianchi [2000] for an overview of attraction phenomena in relative clauses.

(36) Non-standard Russian (Yandex search engine)

²***Vsem***, *komu* *eto* *nado*, *sami* *razberutsja.*
to.everybody to.whom this is.necessary themselves will.figure.out

'Everybody who needs it (lit. to whom it is necessary) will figure it out himself.'

— Case attraction in non-standard Russian [Kholodilova 2015], i.e., constructions in which the relative pronoun receives case from the main clause, as in (37).

(37) Non-standard Russian (Yandex search engine)

Bol'šuju *sobaku* *zvali* *Dunkan,* ***kotoruju*** *pomen'še* — *Tom.*
big.ACC dog.ACC they.called Dunkan which.ACC smaller Tom

'The big dog was called Dunkan, the smaller one was called Tom.'

— Transparent relatives in non-standard Russian [Холодилова 2016]. Transparent relative constructions can be roughly defined as those which allow the nominal predicate of the relative clause to participate in the main clause, for instance, by triggering agreement as in (38), and see [van Riemsdijk 2001] for more details.

(38) Non-standard Russian (RNC)

V *každoj* *strane* ***byli*** *i* *est'* *to,* *čto* *nazyvaetsja* "*social'nye problemy*".
in every country were and is/are that what is.called social problems

'In every country, there are what is called social problems'

Table 9 summarizes the grammatical tendencies these constructions have in common. The list might be non-exhaustive. Fuller data on these tendencies can be found in the works cited above.

Table 9. Russian relative constructions with a grammatical interaction
between the main and the subordinate clause: Common tendencies

		relative pronoun choice	predicate agreement	inverse attraction	case attraction	transparency
"lightness" of the head	no head > pronoun		(logically impossible)	(logically impossible)	+	+
	pronoun > noun	+	+	+	+	+
semantic types	maximalizing > restrictive	+		+	+	
	restrictive > appositive	+	+		+	

Finally, some phenomena suggest that the distinction between more and less integrated heads is typologically relevant.

First, the idea that the opposition between internal and external relative heads is gradual was advanced by Lander [Ландер 2012: 403–417] on the basis of Adyghe and some typological data.

Second, inverse attraction (see the definition above) in Standard Moksha and Standard Udmurt is only possible if the head of the clause does not contain a noun [Kholodilova, Privizentseva 2015], and, according to Aghaei [2003], as cited in Cinque [2015], inverse attraction in Persian is impossible in appositive relative clauses. These facts suggest that inverse attraction tends cross-linguistically to show tendencies similar to those listed in Table 9.

A small piece of evidence in favor of the overall idea is also supplied by Finnish punctuation as related to pausation. If some of the heads are more integrated into the relative clause than others, it is natural to expect that this difference will be reflected in their intonation patterns, which can in turn be reflected by the punctuation. According to [Itkonen, Maamies 2012], a comma can be absent before a relative clause in Finnish if and only if it is headed by a demonstrative pronoun. A corpus study presented in Table 10 shows that a further difference can be found between the singular and the plural of the relative demonstrative pronoun, and commas are left out more often if the head is in the singular. This means that Finnish punctuation obeys at least two of the tendencies described above, the tendency for the demonstrative pronouns to be more integrated than nouns (31) and the tendency for singular to be more liable to integration than plural (33).

Table 10. Commas in Finnish relative clauses with pronominal heads
and the pronoun *joka* 'which' (InterCorp)

	no comma	comma	ratio of missing commas
se 'that'	284	591	0.3
ne 'those'	154	1208	0.1

To summarize, I suggest, following [Ландер 2012: 403–417], that the level of integration of the head into the relative clause is a typologically relevant gradual macro-parameter. I also argue that this macro-parameter includes the opposition between nouns and demonstrative pronouns in the head, as well as that between plural and singular demonstrative pronouns. Finally, I suggest that this macro-parameter manifests itself, among other things, in the choice of relative pronouns in Slavic languages.

Bibliography

Sources

ASPAC,   Amsterdam Slavic Parallel Aligned Corpus compiled by Adrian Barentsen[26]

DOTKO,   Dolnoserbski tekstowy korpus (www.dolnoserbski.de/korpus/, accessed via "Kontext" interface at kontext.korpus.cz)

HOTKO,   Hornolužický textový korpus (korpus.cz/corpora/run.cgi/first_form, accessed via "Kontext" interface at kontext.korpus.cz)

InterCorp (https://ucnk.ff.cuni.cz/intercorp/, accessed via "Kontext" interface at kontext.korpus.cz)

FidaPLUS, Korpus slovenskega jezika

Parasol   (von Waldenfels, Ruprecht and Meyer, Roland (2006–): ParaSol, a Corpus of Slavic and Other Languages. Available at parasolcorpus.org. Bern, Regensburg)

RNC,   Russian National Corpus (ruscorpora.ru), the main and parallel subcorpora

Literature

Aghaei 2003
    Aghaei B., "Case attraction: Evidence for raising analysis for relative clauses in Farsi. Paper presented at SALA 23" (unpublished manuscript).

Bartels, Spiess 2012
    Bartels H., Spiess G., "Restrictive relative clauses in the Sorbian languages," *Sprachtypologie und Universalienforschung*, 65/3, 2012, 221–245.

Bauer 1967
    Bauer J., "K vývoji vztažných vět v slovanských jazycích," *Slavica Slovaca*, 2, 1967, 297–320.

Belyaev, Haug
    Belyaev O., Haug D., "Correlatives in Indo-European and Beyond" (unpublished manuscript).

Bianchi 2000
    Bianchi V., "Some Issues in the Syntax of Relative Determiners," in: *The Syntax of Relative Clauses*, Amsterdam, Philadelphia, 2000, 53–81.

Browne 1986
    Browne W., *Relative Clauses in Serbo-Croatian in Comparison with English*, Zagreb, 1986.

Cinque 2015
    Cinque G., "Three Phenomena Discriminating between "Raising" and "Matching" Relative Clauses," *Semantic-Syntax Interface*, 2, 1, 2015, 1–27.

Citko 2000
    Citko B., "Parallel Merge and the Syntax of Free Relatives" (PhD Thesis, New York, 2000).

---

[26] I am grateful to Adrian A. Barentsen for access to the corpus data.

———— 2004

    Citko B., "On Headed, Headless, and Light-headed Relatives," *Natural Language and Linguistic Theory*, 22/1, 2004, 95–126.

Darovec 1985

    Darovec M., "Vzťažné věty v slovenčině," *Studia Academica Slovaca*, 14, 1985, 107–131.

Dezső 1982

    Dezső L., *Typological Studies in Old Serbo-Croatian Syntax*, Budapest, 1982.

de Vries 2002

    de Vries M., "The Syntax of Relativization" (PhD Thesis, University of Amsterdam, 2002).

Fasske, Michalk 1981

    Fasske H., Michalk S., *Grammatik der obersorbischen Schriftsprache der Gegenwart. Morphologie*, Bautzen, 1981.

Gołąb, Friedman 1972

    Gołąb Z., Friedman V. A., "The *Relative Clause in Slavic*," in: P. Peranteau, J. Levi, G. Phares, eds., *The Chicago Which Hunt. Papers from the Relative Clause Festival*, Chicago, 1972, 30–46.

Grosu, Landman 1998

    Grosu A., Landman F., "Strange Relatives of the Third Kind," *Natural Language Semantics*, 6, 1998, 125–170.

Haspelmath 2010

    Haspelmath M., "Comparative Concepts and Descriptive Categories in Cross-linguistic Studies," *Language*, 86/3, 2010, 663–687.

Heine, Kuteva 2006

    Heine B., Kuteva T., *The Changing Languages of Europe*, Oxford, 2006.

Henderey 2012

    Henderey R., *Relatives in Time and Space: A Case Study in the Methods of Diachronic Typology*, Amsterdam, Philadelphia, 2012.

Itkonen, Maamies 2012

    Itkonen T., Maamies S., *Uusi kieliopas*, 5. painos, Helsinki, 2012.

Janaš 1976

    Janaš P., *Niedersorbische Grammatik für den Gebrauch der Sorbischen Erweiterten Oberschule*, Bautzen, 1976.

Karlík 1988

    Karlík P., "Relativní věty z hlediska reference," *Sborník prací Filozofické fakulty brněnské univerzity. A: Řada jazykovědná*, 37/A36, 59–66.

Kholodilova 2013

    Kholodilova M., "Animacy Distinction in (Mostly) Slavic Possessive Relative Pronouns. Talk at CoQuaT 2013, Leipzig, August 14, 2013" (unpublished manuscript).

———— 2015

    Kholodilova M., "Case Attraction in Non-standard Russian. Talk at 'Insufficient Strength to Defend Its Case': Case Attraction and Related Phenomena, Wrocław, September 18–19, 2015" (unpublished manuscript).

Kholodilova, Privizentseva 2015

    Kholodilova M., Privizentseva M., "Inverse Attraction in Finno-Ugric Languages. Talk at 'Insufficient Strength to Defend Its Case': Case Attraction and Related Phenomena, Wrocław, September 18–19, 2015" (unpublished manuscript).

Kordić 1994
Kordić S., "Relativna rečenica s formalnim antecedentom," *Radovi Zavoda za slavensku filologiju*, 28–29, 1994, 81–93.

——— 1999
Kordić S., *Der Relativsatz im Serbokroatischen*, München, 1999.

Křížková 1970
Křížková H., "Relativní věty v současných slovanských jazycích," *Slávia*, 39, 1970, 10–40.

Lehmann 1984
Lehmann Ch., *Der Relativsatz: Typologie seiner Strukturen; Theorie seiner Funktionen; Kompendium seiner Grammatik*, Tübingen, 1984.

Mithun 2012
Mithun M., "Questionable Relatives," in: B. Comrie, Z. Estrada-Fernández, eds., *Relative Clauses in Languages of the Americas: A Typological Overview*, Amsterdam, Philadelphia, 2012.

Mitrović 2012
Mitrović I., "Relative Clauses in Serbian. Talk at 'Syntax Lab,' September 11, 2012" (unpublished manuscript).

Mucke 1891
Mucke K. E., *Historische und vergleichende Laut- und Formenlehre der niedersorbischen (niederlausitzisch-wendischen) Sprache mit besonderer Berücksichtigung der Grenzdialekte und des Obersorbischen*, Leipzig, 1891.

Murelli 2011
Murelli A., *Relative Constructions in European Non-standard Varieties*, Berlin, Boston, 2011.

Podhajecká 2010
Podhajecká K., "Komunikačný potenciál operátora kto v súčasnej písanej slovenčine," in: *Zborník plných príspevkov z XIX. kolokvia mladých jazykovedcov (Trnava — Modra-Harmónia, 18.–20.11.2009)*, Trnava, 2010, 276–282.

——— 2012
Podhajecká K., "Operátory *čo*, *kto* a *ktorý* ako prostriedky rozvíjania výpovede," in: M. Chovanec, ed., *Studentská vedecká konferencia. Zborník príspevkov*, 7, Prešov, 2012, 757–764.

Polański 1967
Polański K., *Składnia zdania złożonego w języku górnołużyckim*, Wrocław, 1967.

Senderska 2013
Senderska J., "Zdania względne w dobie średniopolskiej," *Polonica*, 33, 2013, 327–342.

Spencer 1993
Spencer D., "Aspects of the Syntax of Relative Clauses in Colloquial and Standard Russian" (PhD Thesis, University of Toronto, 1993).

Svoboda 1967
Svoboda K., "K některým otázkám vět vztažných," *Slavica Slovaca*, 2/4, 1967, 334–339.

Szucsich 2003
Szucsich L., "The Structure of Relative Clauses in Slavic," in: P. Kosta et al., eds., *Investigations into Formal Slavic Linguistics*, 2, Frankfurt a. M., 2003, 697–713.

Vaillant 1958
Vaillant A., *Grammaire comparée des langues slaves*, 2: *Morphologie*, Lyon, 1958.

van Riemsdijk 2001
van Riemsdijk H., "A Far from Simple Matter: Syntactic Reflexes of Syntax-pragmatics Misalignments," in: I. Kenesei, R. M. Harnish, eds., *Semantics, Pragmatics, and Discourse. Perspectives and Connections. A Festschrift for Ferenc Kiefer*, 2001, 21–41.

VON WALDENFELS 2014

VON WALDENFELS R., "Explorations into Variation across Slavic: Taking a Bottom-up Approach," in: B. SZMRECSANYI, B. WÄLCHLI, eds., *Aggregating Dialectology and Typology: Linguistic variation in text and speech, within and across languages*, Berlin, New York, 2014, 290–323.

ZUBATÝ 1918

ZUBATÝ J. "Jenž, který, kdo, co," *Naše řeč*, 2, 1918, 37–44.

БАУЭР 1967

БАУЭР Я., "К развитию относительных придаточных предложений в славянских языках", *Вопросы языкознания*, 5, 1967, 47–59.

БЕЛИЧОВА 1988

БЕЛИЧОВА Е., "Придаточные предложения относительные и посессивность в современных славянских языках", in: Ю. Н. КАРАУЛОВ, ред., *Язык: система и функционирование*, Москва, 1988, 22–30.

БОРКОВСКИЙ 1973

БОРКОВСКИЙ В. И., *Сравнительно-исторический синтаксис восточнославянских языков. Сложноподчиненные предложения*, Москва, 1973.

ЗАЛИЗНЯК, ПАДУЧЕВА 1975

ЗАЛИЗНЯК А. А., ПАДУЧЕВА Е. В., "К типологии относительного предложения", *Семиотика и информатика*, 6, 1975, 51–101.

ИЛИЕВ 2012

ИЛИЕВ И. Г., *Теория на относителността. Българските относителни местоимения — произход и развой*, Пловдив, 2012.

ЛАНДЕР 2012

ЛАНДЕР Ю. А., "Релятивизация в полисинтетическом языке: адыгейские относительные конструкции в типологической перспективе" (Дисс. . . . канд. филол. наук, Москва, Институт востоковедения РАН, 2012).

НИКУНЛАССИ 2008

НИКУНЛАССИ А., *Приместоименно-относительные конструкции в современном русском языке* (= Slavica Helsingiensia, 33), Helsinki, 2008.

НИЦОЛОВА 1986

НИЦОЛОВА Р., *Българските местоимения*, София, 1986.

ПАВЛОВИЋ 2012

ПАВЛОВИЋ С., "Систем релативизатора у српском језику", *Зборник матице српске за славистику*, 82, 2012, 71–86.

ТРУБАЧЕВ 1987

ТРУБАЧЕВ О. Н., ред., *Этимологический словарь славянских языков: праславянский лексический фонд*, 13, Москва, 1987.

ХОЛОДИЛОВА 2015

ХОЛОДИЛОВА М. А., "Согласование с вершиной относительных конструкций и обособленных именных оборотов в русском языке", *Русский язык в научном освещении*, 2, 2015, 74–97.

——— 2016

ХОЛОДИЛОВА М. А., "*(То,) что называется*: между относительной конструкцией и устойчивым вводным оборотом", *Acta linguistica Petropolitana. Труды Института лингвистических исследований РАН*, 12/3, 2016, 763–780.

References

Bartels H., Spiess G., "Restrictive relative clauses in the Sorbian languages," *Sprachtypologie und Universalienforschung*, 65/3, 2012, 221–245.

Bauer J., "On the Development of Relative Subordinate Clauses in the Slavonic Languages," *Voprosy jazykoznanija*, 5, 1967, 47–59.

Bauer J., "K vývoji vztažných vět v slovanských jazycích," *Slavica Slovaca*, 2, 1967, 297–320.

Belichova E., "Pridatochnye predlozheniia otnositel'nye i posessivnost' v sovremennykh slavianskikh iazykakh," in: Yu. N. Karaulov, ed., *Iazyk: sistema i funktsionirovanie*, Moscow, 1988, 22–30.

Bianchi V., "Some Issues in the Syntax of Relative Determiners," in: *The Syntax of Relative Clauses*, Amsterdam, Philadelphia, 2000, 53–81.

Borkovsky V. I., *Sravnitel'no-istoricheskii sintaksis vostochnoslavianskikh iazykov. Slozhnopodchinennye predlozheniia*, Moscow, 1973.

Browne W., *Relative Clauses in Serbo-Croatian in Comparison with English*, Zagreb, 1986.

Cinque G., "Three Phenomena Discriminating between «Raising» and «Matching» Relative Clauses," *Semantic-Syntax Interface*, 2, 1, 2015, 1–27.

Citko B., "On Headed, Headless, and Lightheaded Relatives," *Natural Language and Linguistic Theory*, 22/1, 2004, 95–126.

Darovec M., "Vzťažné věty v slovenčině," *Studia Academica Slovaca*, 14, 1985, 107–131.

Dezső L., *Typological Studies in Old Serbo-Croatian Syntax*, Budapest, 1982.

Faßke H., Michalk S., *Grammatik der obersorbischen Schriftsprache der Gegenwart. Morphologie*, Bautzen, 1981.

Gołąb Z., Friedman V. A., "The Relative Clause in Slavic," in: P. Peranteau, J. Levi, G. Phares, eds., *The Chicago Which Hunt. Papers from the Relative Clause Festival*, Chicago, 1972, 30–46.

Grosu A., Landman F., "Strange Relatives of the Third Kind," *Natural Language Semantics*, 6, 1998, 125–170.

Haspelmath M., "Comparative Concepts and Descriptive Categories in Cross-linguistic Studies," *Language*, 86/3, 2010, 663–687.

Heine B., Kuteva T., *The Changing Languages of Europe*, Oxford, 2006.

Henderey R., *Relatives in Time and Space: A Case Study in the Methods of Diachronic Typology*, Amsterdam, Philadelphia, 2012.

Iliev I. G., *Teoriia na otnositel'nostta. Bŭlgarskite otnositelni mestoimeniia — proizkhod i razvoi*, Plovdiv, 2012.

Itkonen T., Maamies S., *Uusi kieliopas*, 5. painos, Helsinki, 2012.

Janaš P., *Niedersorbische Grammatik für den Gebrauch der Sorbischen Erweiterten Oberschule*, Bautzen, 1976.

Karlík P., "Relativní věty z hlediska reference," *Sborník prací Filozofické fakulty brněnské univerzity. A: Řada jazykovědná*, 37/A36, 59–66.

Kholodilova M. A., "Agreement with the Head of Russian Relative Constructions and Detached Attributive Phrases," *Russkij yazyk v nauchnom osveshchenii (= Russian Language and Linguistic Theory)*, 2, 2015, 74–97.

Kholodilova M. A., "(*To,*) *Čto Nazyvaetsja* '(That) What Is Called': Halfway Between a Relative Construction and a Set Parenthetical Expression," *Acta linguistica Petropolitana*, 12/3, 2016, 763–780.

Kordić S., "Relativna rečenica s formalnim antecedentom," *Radovi Zavoda za slavensku filologiju*, 28–29, 1994, 81–93.

Kordić S., *Der Relativsatz im Serbokroatischen*, München, 1999.

Křížková H., "Relativní věty v současných slovanských jazycích," *Slávia*, 39, 1970, 10–40.

Lehmann Ch., *Der Relativsatz: Typologie seiner Strukturen; Theorie seiner Funktionen; Kompendium seiner Grammatik*, Tübingen, 1984.

Mithun M., "Questionable Relatives," in: B. Comrie, Z. Estrada-Fernández, eds., *Relative Clauses in Languages of the Americas: A Typological Overview*, Amsterdam, Philadelphia, 2012.

Murelli A., *Relative Constructions in European Non-standard Varieties*, Berlin, Boston, 2011.

Nikunlassi A., *Relative Constructions with Pronominal Heads in Contemporary Russian* (= Slavica Helsingiensia, 33), Helsinki, 2008.

Nitsolova R., *Bŭlgarskite mestoimeniia*, Sofia, 1986.

Pavlović S., "Sistem relativizatora u srpskom jeziku," *Zbornik matice srpske za slavistiku*, 82, 2012, 71–86.

Podhajecká K., "Komunikačný potenciál operátora *kto* v súčasnej písanej slovenčine," in: *Zborník plných príspevkov z XIX. kolokvia mladých jazykovedcov (Trnava — Modra-Harmónia, 18.–20.11.2009)*, Trnava, 2010, 276–282.

Podhajecká K., "Operátory *čo*, *kto* a *ktorý* ako prostriedky rozvíjania výpovede," in: M. Chovanec, ed., *Studentská vedecká konferencia. Zborník príspevkov*, 7, Prešov, 2012, 757–764.

Polański K., *Składnia zdania złożonego w języku górnołużyckim*, Wrocław, 1967.

Senderska J., "Zdania względne w dobie średniopolskiej," *Polonica*, 33, 2013, 327–342.

Svoboda K., "K některým otázkám vět vztažných," *Slavica Slovaca*, 2/4, 1967, 334–339.

Szucsich L., "The Structure of Relative Clauses in Slavic," in: P. Kosta et al., eds., *Investigations into Formal Slavic Linguistics*, 2, Frankfurt a. M., 2003, 697–713.

Vaillant A., *Grammaire comparée des langues slaves*, 2: *Morphologie*, Lyon, 1958.

van Riemsdijk H., "A Far from Simple Matter: Syntactic Reflexes of Syntax-pragmatics Misalignments,"

in: I. Kenesei, R. M. Harnish, eds., *Semantics, Pragmatics, and Discourse. Perspectives and Connections. A Festschrift for Ferenc Kiefer*, 2001, 21–41.

von Waldenfels R., "Explorations into Variation across Slavic: Taking a Bottom-up Approach," in: B. Szmrecsanyi, B. Wälchli, eds., *Aggregating Dialectology and Typology: Linguistic variation in text and speech, within and across languages*, Berlin, New York, 2014, 290–323.

Zalizniak A. A., Paducheva E. V., "K tipologii otnositel′nogo predlozheniia," *Semiotics and Informatics*, 6, 1975, 51–101.

Zubatý J. "Jenž, který, kdo, co," *Naše řeč*, 2, 1918, 37–44.

Мария Александровна Холодилова, магистр филологии
Институт лингвистических исследований РАН, младший научный сотрудник Лаборатории типологического изучения языков
199053 С.-Петербург, Тучков пер., 9
Россия/Russia
hol_m@mail.ru