

Animacy Distinction in (Mostly) Slavic Possessive Relative Pronouns

1. Introduction

- The point of departure: in some languages a primarily animate word is used to relativize on inanimate possessors, as in (1).

(1) RUSSIAN, RNC

⟨...⟩ *dal'se pošel tekst, č'e avtorstvo prinadležalo Šeremetu*¹
further went text whose authorship belonged to.Sheremet

‘There followed a text whose author was Sheremet’.

- In what sense is Russian *čej* “primarily animate”?
 - Synchronically: the more frequent interrogative pronoun (2a–b) and indefinite pronouns (2c) refer (almost) exclusively to animates.
 - Diachronically: the interrogative use is in fact etymologically primary, see e.g. (Bauer 1967; Gołąb, Friedman 1972).

(2) RUSSIAN, constructed examples

a. ??*Čej avtor Šeremet?*

whose author Sheremet

lit. ‘Whose author is Sheremet?’

b. ??*Ja ne znaju, čej avtor Šeremet.*

I not know whose author Sheremet

lit. ‘I don’t know whose author Sheremet is’.

c. ??*Čej-to avtor*

whose-INDEF author

lit. ‘someone’s author’

- Cf. a similar distribution in the use of English *whose*: “While the interrogative genitive *whose* is never used of things, the relative *whose* may be used to avoid the postponed *of which*” (Jespersen 2006/1933: 152).

(3) ENGLISH (Jespersen 2006/1933: 152)

A hill, whose peak was still buried in the fog (Stevenson)

- This doesn’t seem to be a usual case with other relativized positions, cf. **The hill who...*
- **Questions** to be partly answered:
 - Where (in which languages)?
 - How often?
 - Why in this very position?
- **Problems** (why make a corpus study):
 - Judgement tasks can be problematic (and, in fact, sometimes are problematic) for the speakers:
 - the construction is “illogical”;
 - there’s often a more frequent and more logical alternative, like *kotorogo* (which.GEN) in Russian;
 - the construction is relatively rare (see below);
 - the construction is often stylistically non-neutral, e. g. “*čijto* is not very common in spoken Bulgarian” (Gołąb, Friedman 1972: 45).

E. g., rejected altogether by some Russian and English speakers, even though the examples are relatively frequent.

¹ Russian, Ukrainian, Belorusan, Bulgarian, Macedonian and Armenian data are transliterated.

- The descriptions seem to be unjustifiably conservative on this point if this issue is mentioned at all, e. g.
 - SerboCroatian: see an overview in Kordić (1995: 130–132);
 - Russian. Švedova (1980: § 2880): *čej* can only be used with reference to humans (+ “personification” + collective nouns in some contexts)

(4) Russian, RNC, Vladimir Orlov. Al’tist Danilov (1980)

⟨...⟩ *bljuda i butyli, č’i aromaty, soki, zapakhi i bukety*
 dishes and bottles whose odours juices, smells and bouquets

kogda-to privlekali junogo Danilova⟨...⟩
 once attracted young Danilov

lit. ‘...dishes and bottles whose odours, smells and bouquets used to attract the young Danilov’

RNC, 1950–2000: about 41 ipm of *čej* with inanimate head (24 with a head clearly unrelated to people)

- Slovene. (Gołąb, Friedman 1972: 41): only personal antecedents for *čigar*.
 Nova Beseda, ~ 1930–now: 7% of inanimate possessor relativization.

→ corpus study, wherever possible

- **Still some problems**

- extremely rare
 - rare position to be relativized, low on the Accessibility Hierarchy, relativization is often avoided, see e.g. (Fiorentino 2007: 270, 279);
 - rare kind of possessor.

→ large corpora needed;

- single examples don’t seem reliable (all the conclusions below are based on multiple examples)

→ even larger corpora and much patience needed.

- Not considered below:
 - free relatives;
 - false free relatives;
 - correlatives.

Probably, pattern with interrogatives (2), but seem most problematic to find and / or evaluate.

2. Slavic *č*-pronouns, the basic data²

- The stem is common Slavic (Vasmer 1986: 323–324).
- Interrogative and indefinite *č*-pronouns invariably refer to animates (Beličova 1988).
- The relative use of interrogative *čej*: late, not Common Slavic (Gołąb, Friedman 1972: 31–32); (Keršiene 1973: 27, 56): Old Russian, Old Ruthenian: *čej* is used only rarely if compared to the modern use.

Table 1. The basic data on the use of *č*-pronouns in headed postnominal relative clauses

		The <i>č</i> -pronoun	some relevant comments in grammatical descriptions	Beličova 1988	The use in headed postnominal relative clauses (with animates and inanimates)		
					ipm (NB: mostly preliminary!)	the size of (sub)corpora	the corpora
West Slavic	Czech	<i>čí</i>		not used in RCs with head nouns	0	1,629,868	ParaSol
	Polish	<i>czyj</i>	(Gołąb, Friedman 1972): OK, but rather colloquial		0	3,395,950	ParaSol
	Slovak	<i>čí</i>			0	1,457,925	ParaSol
	Upper Sorbian	<i>čejiž</i>	(Bartels, Spiess 2012: 226): almost unused		0,3	44,367,372	HOTKO

² I’d like to thank Peter Arkadiev and Philip Minlos for their comments on the Slavic part of the study.

		The č-pronoun	some relevant comments in grammatical descriptions	Beličova 1988	The use in headed postnominal relative clauses (with animates and inanimates)		
					ipm (NB: mostly preliminary!)	the size of (sub)corpora	the corpora
East Slavic	Russian	<i>čej</i>	(Švedova 1980: § 2880): animates (see above)	only used with human reference	42	3,637,357	ParaSol
	Ukrainian	<i>čij</i>			22	1,017,057	ParaSol
	Belorussian	<i>čyj</i>			10	482,467	ParaSol
South Slavic	Slovene	<i>čigar</i>	(Gołąb, Friedman 1972: 41): only personal antecedents	also used with inanimates	9	3,126,259	Nova Beseda, ~ 20th c.
	Bulgarian	<i>čijto</i>					
	Macedonian	<i>čijšto, čij</i>					
	Serbian, Croatian, Bosnian, Montenegrin	<i>čiji</i>					

3. Č-pronouns with inanimate heads: some illustrations

3.1. East Slavic

- RUSSIAN (1)
 - UKRAINIAN
- (5) ParaSol, S. Lem, Golos neba
Vin buv zadumanyj jak golos, čije vidlunnja povernet'sja ⟨...⟩
 it was conceived as voice whose echo will.return
 'It was conceived as a voice whose echo would return...'³.
- BELARUSAN
- (6) Corpus Albaruthenicum, A. I. Lučanok. Udaskanalenne instyucyjnal'naj madeli respubliki Belarus'
 ⟨...⟩ *geagrafičny raen, čye mežy vyznačany*
 geographical area whose boundaries are.defined
jae zdol'nascju abkladac' adatkam paddanykh
 by.its ability impose tax on.citizens
 '...geographical area whose boundaries are defined by its ability [ability of the state] to tax the citizens'.

3.2. South Slavic

- BULGARIAN
- (7) Intercorp
 ⟨...⟩ *produkti, čijto naj-goljam tegloven komponent e juta*
 products whose most heavy component is jute
 '...products whose largest component by weight is jute'
- MACEDONIAN
- (8) ParaSol, U. Eko. Imeto na rozata
 ⟨...⟩ *Firenca za čii prekrasni crkvi imavme slušnato nogu falbi*
 Florence about whose beautiful churches we.had heard many praises
 '...Florence, whose churches I had heard praised as most beautiful'
- SLOVENE
- (9) Nova beseda, Damijan Šinigoj. Neizstreljeni naboj
metek ⟨...⟩ *čigar inicialna kapica je bila le nekaj desetink limetra*
 bullet whose capsule had been only some tenths of.millimeter

³ Where possible, the original / professional translations are given.

oddaljena od udarne igle
distanced from firing pin

- SERBIAN (10), CROATIAN (11), BOSNIAN (12), MONTENEGRIN (13)
- (10) RNC, D. Braun. Da Vinčijev kod
 ⟨...⟩ *u ovakvim zgradama, čiji su¹ stanovnici provodili² veći deo dana*
 in these buildings whose inhabitants spent^{1,2} larger part of.day
u molitvi, u svojim sobama
 in prayer in their rooms
 ‘...in these halls where the residents spent most of the day in their rooms in prayer’
- (11) Intercorp, M. Kundera. Besmrtnost
 ⟨...⟩ *prije svadbe (u čiju je¹ nužnost ipak u dubini duše mnjao²)*
 before wedding in whose necessity still in depth of.soul he.doubted^{1,2}
 ‘...before his wedding day (whose necessity he doubted in his heart of hearts)’
- (12) The Oslo Corpus of Bosnian Texts
 ⟨...⟩ *preko drvene ograde na čijem je¹ vrhu bila² bodljikava žica*
 over wooden fence on whose top was^{1,2} barbed wire
 ‘...over a wooden fence with barbed wire on top’
- (13) Montekorpus
Predlog zakona čiji predlagač nije Vlada ⟨...⟩
 proposal of.law whose proposer is.not Government
 ‘...a proposed law whose proposer is not Government...’

4. Some idea of frequencies

- Yandex search engine;
- Why do it:
 - to make sure the relevant examples are
 - rather numerous;
 - not just calqued (as could be the case in translational corpora);
 - not outdated.
 - to get some idea of the difference between the languages in relatively comparable sets of data;
- Search conditions:
 - only masculine singular heads (\approx the distribution of Slovene *čigar* [Derbyshire 1993: 52]);
 - only the relative clauses with the possessive phrase on the first place (not even a preposition before);
 - only the relative clauses with single-word possessives.
 - the last (chronologically) sentences with one of the relativizers in Yandex (no more than 100);

Table 2. The frequency of *č*-pronouns as compared to other means of internal possessor relativization (Yandex)

	East Slavic			South Slavic			
	Russian (<i>čej,</i> <i>kotorogo</i>)	Ukrainian (<i>čij,</i> <i>jakogo</i>)	Belarusan (<i>čyj,</i> <i>jakoga</i>)	Slovene (<i>čigar,</i> <i>katerega</i>)	Croatian (<i>čiji,</i> <i>kojeg(a)</i>)	Bulgarian (<i>čijto, –</i>)	Macedonian (<i>čijšto, čij, –</i>)
‘man’ (‘human’)	0,59 (59/100)	0,83 (83/100)	0,49 (49/100)	0,49 (42/85)	0,99 (99/100)	(1)	(1)
‘man’ (‘male’)	0,43 (43/100)	0,45 (45/100)	0,6 (9/25)	0,48 (15/31)	1,00 (100/100)	(1)	(1)
‘institute’	0,09 (9/100)	0,08 (8/100)	0,11 (1/9)	0,00 (0/10)	0,98 (44/45)	(1)	(1)
‘film’	0,03 (3/100)	0,03 (3/100)	0,00 (0/38)	0,05 (5/100)	0,99 (99/100)	(1)	(1)

Comments on the table:

- Slovene: patterns with East Slavic;
 - 18–21 c.: a decrease in the use of *čigar* with animate heads, an even more dramatic decrease in the use of *čigar* with inanimate heads;

Table 3. Diachronical data on the frequency of Slovene *čigar* (Nova Beseda)

	The author's year of birth		
	1751–1850	1851–1900	1901–2000
1. ipm with animate heads (<i>čigar</i>)	48 (69)	29 (136)	9 (29)
2. ratio of <i>čigar</i> with animate heads (<i>čigar</i> , <i>katerega</i>)	0,9 (69/75)	0,9 (136/151)	0,7 (29/44)
3. ratio of <i>čigar</i> with inanimate heads (<i>čigar</i> , <i>katerega</i>)	0,3 (8/23)	0,4 (26/64)	0,1 (4/57)
4. 2./3.	2,7	2,4	9,5

Could be indirectly related to the fact Slovene, unlike other South Slavic languages, uses the relative pronoun *kateri*, which is cognate with Russian *kotoryj*.

- East Slavic and Slovene: at least some examples of *č*-pronouns with inanimate heads in non-translations;
- East Slavic and Slovene: still a strong correlation between animacy and the choice of *č*-pronouns.

5. More data

**(My apologies: not as corpus-based,
not as quantitative,
not at all Slavic,
work deep in progress)**

- Really few languages which could be relevant:
 - a very limited set of languages with relative pronouns, a well-known SAE feature, see (Haspelmath 2001: 1494–1495)
 - not just any language with relative pronouns:
 - in some – no animacy distinction in relative pronouns (e.g. Latvian);
 - in some – the “animate” relative pronoun is never or almost never used with nominal heads (e.g. Finnish);
 - some: not enough data available (e. g., Romani).
 - ENGLISH (3)
 - DUTCH
- (14) InterCorp, Douglas Adams. Het Transgalactisch Liftershandboek
Een computer wiens operationele parameters ik nog niet waard
 a computer who.GEN.M operational parameters I yet not worth
ben uit te rekenen
 am out to calculate
 ‘A computer whose merest operational parameters I am not worthy to calculate’
 Google: “huis waarvan de eigenaar” OR “huis wiens eigenaar” (house whose owner) – 48 : 4, not frequent, but probably, not accidental

- INGRIAN FINNISH

(15) GEN, possessor: **shift**

nurka-s seisò kuappi kenen / ?minkän ukse-t o-vat kīn
 corner-IN stay.3SG cupboard who.GEN what.GEN door-PL e-3PL closed
 ‘In the corner, there’s a cupboard with its doors closed’.

(16) NOM, SU: **no shift**

*poütä mikä / *kuka seisò nurka-s on matala*
 table what who stay.3SG corner-IN be.3SG low
 ‘The table in the corner is low’.

(17) GEN, in PP: **no shift**

*nurka-s seisò kuappi minkän / *kenen piäl seisò lamppu*
 corner-IN stay.3SG cupboard what.GEN who.GEN on stay.3SG lamp
 ‘In the corner, there’s a cupboard with a lamp on it’.

(18) GEN, DO: **no shift**

*nurka-s seisò kuappi minkän / *kenen miä ost-i-n eilen*
 corner-IN stay.3SG cupboard what.GEN / who.GEN I buy-PST-1SG yesterday
 ‘In the corner, there’s a cupboard I bought yesterday’.

▪ HUNGARIAN?

- controversial data;
- generally rejected by the speakers or accepted as only marginally possible⁴, but...
- ...relatively frequent on the net:

(19) Blikk.hu, 2011 (Hungarian daily)

Tizsakécske és Lakitelek között volt egy nádtetős ház, aki-nek a tulajdonosa
 Tizsakécske and Lakitelek among was one thatched house who-DAT the owner
nem egyezett bele a bontásba <...>
 not agreed into the wrecking-ILL

‘Between Tizsakécske and Lakitelek, there was a thatched house, whose owner didn’t give permission for its demolition’.

- could be some typos...



- ...however, they would be unbelievably frequent exactly in the expected positions:

Table 4. The frequency of Hungarian animate relative pronouns in different kinds of relative clauses, as attested in Google

				<i>aki-</i>	<i>ami-</i>	
<i>ház</i>	<i>aminek</i>	/	<i>akinek</i>	<i>a tulajdonosa</i>	5	11
house	what.GEN		who.GEN	owner		
<i>ház</i>	<i>amiben</i>	/	<i>akiben</i>		0	82
house	in.which		in.who			
9.02.13–9.03.13						

Fischer’s exact test, $p = 0,00006 < 0.01$.

▪ ARMENIAN⁵

(20) constructed

OKšat em¹ sir-um² ays tek’st-ə, um heyhinak-ə inj,
 very like-PRS.1SG^{1,2} this text[NOM]-DEF who.GEN author[NOM]-DEF I.DAT
c’avok’, haytni če
 unfortunately unknown NEG.COP.PRS.3SG
 lit. ‘I really like this text, whose author is unknown to me’.

⁴ I wish to thank Klara Malova, Dorottya Dolovai, Virág Vajda and Szilvia Turoczki for their judgments. I am also grateful to Egor Kashkin for his help in getting in touch with Hungarian speakers.

⁵ Many thanks to Vasilisa Krylova for constructing this sentence and checking it with Armenian native speakers.

- To sum up, non-Slavic languages with some kind of relatively frequent animate relative pronouns used for possessor relativization: is there a shift in animacy?
 - **Yes:**
 - English;
 - Dutch;
 - Ingrian Finnish;
 - Armenian.
 - **Yes?**
 - Hungarian.
 - Probably, **no** (I did my best to find an example in large corpora and Google, but I didn't):
 - Estonian.
 - **Not (yet?) enough data:**
 - Mordvin;
 - Romani;
 - ...
 - some other, but really few, at least in the sample of (Haspelmath 2001).
- There seems to be a strong tendency for the animacy shift to happen.

6. Why? What is so special about (in)animate possessors?

- Reasons which could “work” within some of the languages (NB: some of them are not actual explanations, but rather observations that could be relevant):
 - 1. Language-internal structural reasons:
 - 1.1. Some languages have no other strategy to relativize on internal inanimate possessors or its use is restricted, cf. “...the relative *whose* may be used to avoid the postponed *of which*” (Jespersen 2006 (1933): 152).
 - 1.2. In some languages, the competing strategies differ in the position of the relativizer. It might be more natural for a possessive relativizer to come first. Possessors are likely to be used for “anchoring”, i. e. introducing referents into the discourse (Koptjevskaja-Tamm 2002).

(21) (Paduceva 2004: 355): “Principle of referential compactness:

A referentially autonomous NP (in particular, a NP the reference of which is known to the speakers) must be introduced into the denotative space of the utterance earlier than its referentially dependent NPs”.

- 1.3. In some languages, the strategies differ in the presence of agreement between the possessee and the relative pronoun.
 - ~ (Beličova 1988): Czech *jejíž* (which.GEN.F) has developed a paradigm (*jejího*, *jejímu*).
- 1.4. In some languages, the strategies differ in the presence of (independent) case marking on the relative pronoun.
- 1.5. In some languages the “shifting” pronoun is more “specialized” in that it can only be used for possessor relativization.
 - 1.4, 1.5. ~ “Light” “specialized” markers of different origin used for possessor relativization in French (*dont*), Swedish (*vars*), Spanish (*cuyo*).
- 2. Contact-**induced** changes
 - (Hendery 2012: 235–237): a major factor in relative clause changes;
 - (Fiorentino 2007): much in common between the systems of relativizers in European languages;
 - (Issatschenko 1983: 514): The relative use of Russian *čej* was, probably, influenced by French *dont*;
 - Ingrian Finnish: massive Russian influence;
 - Slavic languages: not only genetic links, but also much interference;
 - other languages: you never know...

Table 5. Available explanations of the pattern in different languages

	1.1. (almost) no choice	1.2. difference in word order	1.3. difference in +/- agreement	1.4. difference in +/- case	1.5. a more specialized marker	2. contact influence expected
Russian	–	+ ⁶	+	+	+	? (French)
Ukrainian	–	+	+	+	+	? (other Slavic)
Belarusian	–	+	+	+	+	? (other Slavic)
Slovene	–	–	–	+	+	? (other Slavic)
Serbian, Croatian, Bosnian, Montenegrin	–	+/-	+	+	+	? (other Slavic)
Bulgarian	+					? (other Slavic)
Macedonian	+					? (other Slavic)
English	%	+	–	+	+	?
Dutch	–	–	+	–	+	?
Armenian	–	–	–	–	–	?
Ingrian Finnish	%	–	–	–	–	+ (Russian)
? Hungarian	–	–	–	–	–	?

Ergo: The language-internal structural reasons (any of them or all of them) might be relevant, but it's not the case that one of them determines the shift. Thus, the “trigger” is either none of them or a combination thereof.

One more explanation to be considered (however difficult to check):

- In possessors, there seems to be more interplay between animacy, definiteness and the kind of relation between the possessor and the possessee, partly along the lines of the Animacy Hierarchy (Silverstein 1976).
- Probably, that is because possessors are not actual participants of the situation and what matters is not the kind of object but rather its relation to the discourse.

-
- E.g., the distinction between proper/common nouns is partly independent from animacy:
 - Slovene, inanimate nouns, χ^2 , $p \approx 0,01$

Table 6. The frequency of animate proper nouns as heads of Slovene relative clauses with *čigar* and *kateri* (Nova Beseda)

	<i>čigar</i> (whose)	<i>katerega</i> (which.GEN)	% of <i>čigar</i>
animate proper nouns	69	4	95 %
animate common nouns	165	34	83 %

- Russian, inanimate nouns, χ^2 , $p < 0,01$.

Table 7. The frequency of inanimate proper nouns as heads of relative clauses with *čej* and *kotorij* (RNC)

	<i>čej</i> (whose)	<i>kotorogo</i> (which.GEN)	% of <i>čej</i>
inanimate proper nouns	113	1479	7,0 %
inanimate common nouns	777	15100	4,9 %

⁶ Though not in the 18th century when these changes started, for the details see (Kholodilova 2011: 12–16).

References

- Bartels H., Speiss G. Restrictive relative clauses in the Sorbian languages. In *STUF, Akademie Verlag* 65, 3, 2012. P. 221–245.
- Bauer Ja. K razvitiju odnositel'nykh pridatočnykh predloženij v slavjanskikh jazykakh. In *Voprosy jazykoznanija* 5, 1967. P. 47–59.
- Beličova E. Pridatočne predloženija odnositel'nye i possessivnost' v sovremennykh slavjanskikh jazykakh. In Karaulov Ju. N. (ed.). *Jazyk: sistema i funkcionirovanie*. Moscow, 1988. P. 22–30.
- Derbyshire W. W. *A Basic Reference Grammar of Slovene*. Columbus, Ohio: Slavica Publishers, 1993.
- Fiorentino G. European relative clauses and the uniqueness of the Relative Pronoun Type. In *Rivista di Linguistica* 19, 2, 2007. P. 263–291.
- Gołąb Z., Friedman V. A. The relative clause in Slavic. In Peranteau P., Levi J., & Phares G. (eds.). *The Chicago Which Hunt. Papers from the Relative Clause Festival*. Chicago, 1972. P. 30–46.
- Haspelmath M. The European linguistic area: Standard Average European. In Haspelmath M. et al. (ed.). *Language Typology and Language Universals. An International Handbook*. Berlin/New York: Walter de Gruyter, 2001. P. 1492–1510.
- Henderey R. *Relative clauses in time and space: a case study in the methods of diachronic typology*. Amsterdam/Philadelphia: John Benjamins, 2012.
- Issatschenko A. *Geschichte der russischen Sprache*. Bd. 2. 1983.
- Jespersen O. *Essentials of English Grammar*. London: Routledge, 2006. — First published 1933.
- Keršiene R. B. Složnopodčinennye opredelitel'nye predloženija. In Borkovskij V. I. (ed.). *Sravnitel'no-istoričeskij sintaksis vostočnoslavjanskikh jazykov. Složnopodčinennye predloženija*. M.: Nauka, 1973.
- Kholodilova M. A. Reljativizacija pozicij posessora v ruskom jazyke. In *Russkij jazyk v naučnom osveščeenii* 21, 2011. P. 5–46.
- Koptjevskaja-Tamm M. Adnominal possession in the European languages: form and function. In *Sprachtypologie und Universalienforschung*, 55, 2. 2002. P. 141–172.
- Kordić S. *Relativna rečenica*. Zagreb, 1995.
- Paducheva E. V. Splitting of Possessive NPs and External Possessor in Russian. In Kim J., Lander Yu. A., Partee B. H. (eds.). *Possessives and beyond: Semantics and Syntax*. Amherst (Mass.): GLSA. 2004: 351–363.
- Silverstein M. Hierarchy of features and ergativity. In Dixon R. M. W. (ed.). *Grammatical categories in Australian languages*. New-York: Humanities press, 1976. P. 112–171.
- Švedova N. Ju. (ed.). *Russkaja grammatika*. V. II. Moscow, 1980.
- Vasmer M. *Etimologičeskij slovar' russkogo jazyka*. V. 4. Moscow, 1986.

The corpora:

- The Oslo Corpus of Bosnian Texts (tekstlab.uio.no/Bosnian/Corpus.html)
- Corpus Albaruthenicum (grid.bntu.by/corpus)
- HOTKO – Hornolužický textový korpus (korpus.cz/corpora/run.cgi/first_form);
- InterCorp (www.korpus.cz)
- Montekorpus (eiprevod.gov.me/korpus)
- Nova beseda (bos.zrc-sazu.si/s_beseda.html)
- ParaSol – A Parallel Corpus of Slavic and other languages (parasol.unibe.ch)
- RNC – Russian National Corpus (ruscorpora.ru)
- Yandex – the Yandex search engine (yandex.ru)