

## Headed Correlatives in Russian\*

### 1 Correlatives: An Introduction

- (1) [ Kogo ljublju ], [ tomu darju ].  
who.ACC love.1SG that.DAT.SG give.away.1SG  
[<sub>RC</sub> wh- ] [<sub>MC</sub> correlate ]
- (2) Kak rabotal, tak i zarabotal.  
how worked that.way PRT earned

For Lehmann (1986), correlatives are a “variety of the adjoined relative construction” (i.e. a construction where a RC is situated outside the (rest of) MC). In fact some analyses claim that RC originates within MC (see below).

Is headedness compatible with correlativity? In other words, how about e.g.

- (3) [ Tomu, [ kogo ljublju ] ], [ tomu darju ]?

- ⊖ Vries (2002, 40): lack of external head as one of the defining properties of correlatives, i.e. HCRs ruled out by definition (which was apparently empirically inspired though). Some deny the presence of the DP/NP level in correlative RCs (Lipták 2009, 8).
- ⊕ Mahajan (2000, 214–215): Hindi features correlative clauses headed by a determiner. Mahajan gives examples with demonstratives (4); note the presence of the NP *a:dmi:*. Similar Indo-Aryan examples, including those with a full NP both in RC and MC, are discussed as somewhat marginal in Cinque (2010).
- (4) vo a:dmi: jo si:ta:ko acc<sup>h</sup>a: lagta: he mujhe vo pasand nahī: he  
DEM man REL Sita.DAT nice seem.IMP be.PRES I.DAT DEM like not be.PRES  
'I do not like the man who Sita likes'

#### In the talk:

- the variety of headed correlatives (HCRs) in Russian
- preliminary diachronic/lexical data; some implications for theories of correlatives

### 2 HCRs in Russian

First mention known to me: Лютикова (2008); also Холодилова (2010). Cf. Tsedryk (n.d., fn. 11), apparently unaware of predecessors:

The anonymous reviewer points out ... (i), suggesting that correlative CPs could involve covert universal quantification and bringing up a possibility of analysis in terms of QR (along the lines of antecedent contained deletion).

- (i) Vsě, čto ty zaxočeš', to ja i kuplju.  
all what.ACC you will.want that.ACC I EMPH will.buy  
[Lit.: 'What you want, that (thing) I'll buy.']

However, the reviewer's example in (i) appears to be an argument against CP movement in terms of universal QR. If (i) is to be derived by moving *vsě, čto ty zaxočeš'* from Dem, we expect this string of words to be a constituent at the base contrary to fact.

\*Thanks to Maria Kholodilova, Olga Mitrenina and Sergey Say for their comments and suggestions.

They are robustly attested in corpus data:<sup>1</sup>

- (5) [ Tot, kto prestupaet ], tot i s bogom prestupal,  
 that.NOM.SG who.NOM trespass.3SG that.NOM.SG PRT with God.INS trespassed  
 poklony bil i vsě ravno prestupal. [Даниил Гранин. Зубр (1987)]  
 bowed and nevertheless trespassed
- (6) [ Vsě, čto polezno čelovečstvu ], to i  
 everything.NOM that.NOM.SG useful.SG.N humankind.DAT that.NOM.SG PRT  
 blagorodno! [Ф. М. Достоевский. Преступление и наказание (1866)]  
 noble.SG.N

Checklist for correlatives (Vries 2002, 40):

- The head is internal. ✓
- The semantics is maximalizing. ✓ (sometimes overt quantification)
- Left-adjoined to the matrix clause. ✓
- Matrix contains a personal or demonstrative pronoun (the correlate) that refers to the modified relative head. ✓ (pronoun sometimes modified)
- They are not nominalized (i.e. not DPs), but they are bare sentences.
  - Do not occur in DP positions. ✓
  - Never have an external determiner. ✗ (attributive demonstrative or quantifier + [<sub>NP</sub> ∅])
  - Never have an external Case ending or another nominal marking. ✓
  - Never have an external adposition. ✓

For *potomu čto* ‘because’ the headless correlative configuration is dubious:

- (7) ?Počemu èto odejanie nadel Ivan, potomu (že) ego nadel (i) Pètr.  
 why this attire.ACC.SG put.on I. that’s.why PRT it.ACC put.on PRT P.

The corresponding HCR-like construction is as in

- (8) (Začem že on nadel čužoe odejan’e? ‘Why on earth did he put on another’s attire?’)  
 Potomu čto lučše, potomu i nadel... [Н. В. Гоголь. Тарас Бульба (1835-1841)]  
 because better that’s.why PRT put.on

Both demonstrative and universal heads are **light heads** in the sense of Citko (2004), which have the properties intermediate between those of regular heads and the absence of a head (see e.g. Kholodilova 2017, 133), e.g. in terms of the admissible relative pronouns (‘what’ vs. ‘which’) in the languages that have restrictions here.

## 2.1 Real Heads

The head is predominantly nominative (if case-marked): among the case-marked heads in the data in Section 3, 30 are NOM, 3 are ACC, homonymous with NOM, and 1 is INS. May it be that the head in fact forms an unanalysable complex with the relativiser (as proposed for the complementiser *to čto* in Korotaev 2013)? No:

<sup>1</sup>Data from the Russian National Corpus, <http://ruscorpora.ru>.

- the head may inflect for case, number and gender  
 (9) U tex, komu platjat, u tex i sprašivajte. [Google]  
 PREP that.GEN.PL who.DAT pay.3PL PREP that.GEN.PL PRT ask.IMPER.PL
- the head may be separated from the *wh*-word, at least with a preposition:  
 (10) — K tomu, k čemu nado, k tomu i  
 to that.DAT.SG to what.DAT should.be to that.DAT.SG PRT  
 odnositsja! [K. Г. Паустовский. Золотая роза (1955)]  
 relate.3SG

## 2.2 Not Topics

A possible objection: fronted clauses in (5)–(6) are just topics; cf. with a DP:

- (11) Lěn'ka navernjaka ne znaet ničego, uroven' ne tot, a vot  
 L. surely NEG know.3SG nothing level.NOM.SG NEG that.NOM.SG but PRT  
 Iljuša ↗ — tot znaet. [Александра Маринина. Последний рассвет (2013)]  
 I. that.NOM.SG know.3SG
- (12) “JUKOSu” ↗ — tomu voobšče ot publičnogo obsuždenija prjamaja  
 YUKOS.DAT that.DAT.SG at.all from public discussion.GEN.SG direct  
 vygoda. [Александр Добровинский. Заказ избирателей (2003) // «Новая газета»]  
 profit.NOM.SG

Rejoinders:

- topicalisation has distinctive **intonation**, which is (in my judgement) not required by (5)–(6)
- universal DPs don't topicalise easily (13), but HCRs with universal heads are OK (14)  
 (13) \*Každyj (čelovek) — tot i ezdit.  
 every man.NOM.SG that.NOM.SG PRT ride.3SG
- (14) Ved' na nej každyj, kto zaxočet, tot i  
 PRT on she.LOC everyone.NOM who.NOM want.FUT.3SG that.NOM.SG PRT  
 ezdit! [Л. Н. Разумовская. Счастье (1981)]  
 ride.3SG ~ 'whoever wants to, exploits her'
- demonstrative DPs are degraded as topics “resumed” by a demonstrative (15), which does not preclude HRCs with demonstratives as correlates  
 (15) ?Tot čelovek — tot na nej i ezdit.  
 that man.NOM.SG that.NOM.SG on she.LOC PRT ride.3SG

On the other hand, the length of the fronted constituent may play a role: the longer it is, the better the demonstrative.

Headed RCs to which correspond personal pronouns may indeed be topics (the semantics is referential, not maximalising):

- (16) Tol'ko ved' te, kogo poslali, — oni tam stesnjat'sja ne  
 just PRT that.NOM.PL who.ACC sent they.NOM there feel.shy.INF NEG  
 stanut. [Google]  
 be.FUT.3PL

### 3 Diachronic and Word-Specific Patterns

**Warning!** The data are quite preliminary. I consider only correlates immediately followed by the particle *i*, which is in fact optional (17) and is not tightly linked to the correlate in terms of linear position (18).

The rest of the main clause in a correlative is often preceded by *i* emphatic particle that marks the most salient part of the sentence.

(Mitrenina 2010, referring to Šimčuk and Ščur)

- (17) Poèтому každyj, kto potruditsja, tot tuda i  
 that's.wh.y everyone.NOM who.NOM work.FUT.3SG that.NOM.SG there PRT  
 vxodit, a kto leniv na duxovnoe, tot idët v preispodnjuju, v  
 enter.3SG but who.NOM lazy spiritually that.NOM go.3SG to hell.ACC to  
 geennu. [Протоиерей Димитрий Смирнов. Проповеди (1984-1989)]  
 Gehenna.ACC
- (18) Každyj, kto bljaxu nadel, tot **nami** i  
 everyone.NOM who.NOM badge.ACC.SG put.on that.NOM.SG we.INS PRT  
 romykaet. [А. А. Богданов. Потапыч (1897-1899)]  
 rule.3SG

To put a limit on search results, the distance between *wh*- and correlate was set to 1–5.<sup>2</sup>

#### 3.1 Lexical Idiosyncrasies

	% headed	<i>skol'ko</i>	<i>kak</i>	<i>kogda</i>	<i>kakoj</i>	<i>kto</i>	<i>čto</i>
<i>gde</i>	4,81	***	***	***	**	***	—
<i>čto</i>	3,54	***	**	**	*	*	
<i>kto</i>	1,66	*	(*)	— <sup>3</sup>	—		
<i>kakoj</i>	0,66	—	—	—			
<i>kogda</i>	0,29	—	—				
<i>kak</i>	0,24	—					

Table 1: Differences in the proportions of headed uses in the main subcorpus of RNC (specifications from Table 2 apply); two-sided  $\chi^2$ ; \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$

- Two poles: *gde* and *čto* vs. *kakoj*, *kogda* and *kak*, with *kto* in-between.

The demonstrative heads corresponding to *gde* and *čto* are monosyllabic, may be closer to the prototype of “lightness”. The same holds for *kak*, but here a considerable proportion of uses are **not** maximalising but rather convey the meaning of ‘just in the same way’.

#### 3.2 Diachronic Variation

See Table 2.

- The proportion of non-demonstrative (universal) heads decreases over time.
- The proportion of demonstrative heads increases.
- Hence, weak overall tendencies.

<sup>2</sup>Another warning: Холодилова (2010) comes to different conclusions regarding diachrony.

<sup>3</sup>Two-sided exact Fisher test:  $p \simeq 0,06$ .

Wh-word	Period	DEM head	Other heads	All heads	Headless	All	Significance
gde 'where'	-1917	5	3	8	239	247	* for DEM, ** for others
	1918-	34	0	34	592	626	
	overall	39	3	42	831	873	
kak 'how' <sup>a</sup>	-1917	1	0	1	115	116	
	1918-	0	0	0	303	303	
	overall	1	0	1	418	419	
kakoj 'which (kind of)'	-1917	1 <sup>b</sup>	0	1	64	65	
	1918-	1	0	1	239	240	
	overall	2	0	2	303	305	
kogda 'when'	-1917	0	0	0	107	107	
	1918-	0	1	1	232	233	
	overall	0	1	1	339	340	
kto 'who'	-1917	1	1	2	398	400	* for all and for DEM
	1918-	15	2	17	730	747	
	overall	16	3	19	1128	1147	
skol'ko 'how many/much'	-1917	0	0	0	120	120	
	1918-	0	0	0	295	295	
	overall	0	0	0	415	415	
čto 'what' <sup>c</sup>	-1917	4	3	7	187	194	
	1918-	8	2	10	276	286	
	overall	12	5	17	463	480	
all	-1917	12	7	19	1230	1249	** for DEM, * for others
	1918-	58	5	63	2667	2730	
	overall	70	12	82	3897	3979	

Table 2: Headed and headless CRs in the main subcorpus of RNC (specified further where necessary); two-sided  $\chi^2$  where applicable

<sup>a</sup>Partial sample: only words starting with A or B after i. Excluded: all uses of kak... tak i conjoining anything but finite verb phrases.

<sup>b</sup>Here the head is heavy, but mind the fact that kakoj is attributive: ...i s takim že licom, s kakim prišla, s takim i tušla 'and she went out with the same face as when she came in' [A. Ф. Писемский. Вспоминаемое море (1863)].

<sup>c</sup>Partial sample: only words starting with A-I after i. Excluded: čto kasajesja 'as regards', čem... tem 'the more... the more'. One pre-1917 example with čto 'instead of to as correlate.

A universal head may have its semantic contribution, removing **homogeneity** (Križ 2015): there is a sense in which (19a) is neither decisively true nor decisively false in Scenario 1, but (19b) is definitely false.

- (19) a. The professors smiled. ibid., v, 11  
 b. All the professors smiled.

**Scenario 1.** *There were ten professors, five of them smiled at a joke.*

Adding a demonstrative head, unlike with universal head, makes no semantic contribution. Hypothesis: headedness becomes more automatic, perhaps with some degree of fixedness. Given the emergent use of *to čto* instead of *čto* with mental and speech verbs (Korotaev 2013; Богданова-Бегларян 2015), this suggests a tendency toward formal headedness in contemporary Russian subordinate clauses.

## 4 Theoretical Implications

1. Some theorists (Mahajan 2000; Mitrenina 2010; Pietraszko 2015) argue that the correlative clause undergoes movement from the clause containing the correlate.

- (A) [<sub>RC</sub> *kto ...*]<sub>i</sub> ... [<sub>MC</sub> [<sub>tot</sub> [<sub>RC</sub> *kto ...*]]<sub>i</sub> ...]

Here it is unclear how the head of the RC emerges.

- (B) [<sub>tot</sub> [<sub>RC</sub> *kto ...*]]<sub>i</sub> ... [<sub>MC</sub> [<sub>tot</sub> [<sub>RC</sub> *kto ...*]]<sub>i</sub> ...] (incomplete deletion, Mahajan 2000)

Option (B) fares well insofar as the head of the HCR is identical to the correlate. However, with universal heads (6) there is no identity. To save the movement account, one would have to stipulate that *to* and *vsě* are different realisations of the same lexeme, cf. Fox, Johnson 2016 for *the* and *every*. (This problem was apparently unknown to Лютикова (2008), but Холодиллова (2010) reports several kinds of mismatch between the correlate and the RC head.)

But even this may be insufficient, as for some speakers (prominently for Leo Tolstoy) the complex ‘ $\forall$  + DEM’ is available as head (20), but not as correlate.

- (20) *Vsě to, čto daët mne Otec, to all that.NOM.SG what.NOM give.3SG I.DAT Father.NOM.SG that.NOM.SG prijdet ko mne... [Л. Н. Толстой. Соединение и перевод... (1902)] come.FUT.3SG to I.DAT*

2. Headless CRs can be “multi-head” (Bhatt 2003), as in (21) from Митренина (2008):

- (21) *A vpročem, ja takogo mnjenja: komu<sub>1</sub> kakaja<sub>2</sub> premudrost’ dalas’, but actually I such opinion.GEN.SG who.DAT which wisdom.NOM.SG was.given tot<sub>1</sub> toj<sub>2</sub> i priderživajsja! [И. С. Тургенев. Собака (1847-1852)] that.NOM.SG that.GEN.SG PRT hold.on.to.IMPER*

In HCRs, the use of relative pronouns not connected to the head is degraded. Tsedryk (n.d.): bare interrogative pronouns with indefinite interpretation are licensed in the (immediate) scope of a quantifier over alternatives, e.g. *esli* ‘if’ (Yanovich 2005). However, Tsedryk discards the idea that (headless) CRs are conditional (it cannot get the movement right).

- (?) But should we draw a divide between headless RCs and HRCs on the grounds of (21)?

## References

- Bhatt R.* Locality in correlatives // *Natural Language & Linguistic Theory*. — 2003. — Vol. 21, no. 3. — Pp. 485–541.
- Cinque G.* Five notes on correlatives // *University of Venice Working Papers in Linguistics*. — 2010. — Vol. 19. — Pp. 35–60.
- Čitko B.* On headed, headless, and light-headed relatives // *Natural Language & Linguistic Theory*. — 2004. — Vol. 22, no. 1. — Pp. 95–126.
- Fox D., Johnson K.* QR Is Restrictor Sharing // *Proceedings of WCCFL*. Vol. 33. — 2016. — Pp. 1–16.
- Kholodilova M.* Competition between ‘Who’ and ‘Which’ in Slavic Light-Headed Relative Clauses // *Slověne*. — 2017. — No. 1. — Pp. 118–147.
- Korotaev N. A.* Clause-combining with *to chto* in Russian spoken discourse // *Computational Linguistics and Intellectual Technologies: papers from the Annual conference “Dialogue”*. — 2013. — Pp. 358–367.
- Križ M.* Aspects of homogeneity in the semantics of natural language: PhD thesis / Križ Manuel. — University of Vienna, 2015.
- Lehmann C.* On the typology of relative clauses // *Linguistics*. — 1986. — Vol. 24, no. 4. — Pp. 663–680.
- Lipták A.* The landscape of correlatives: An empirical and analytical survey // *Correlatives cross-linguistically*. — John Benjamins, 2009. — Pp. 1–48.
- Mahajan A.* Relative asymmetries and Hindi correlatives // *The syntax of relative clauses*. — John Benjamins, 2000. — Pp. 201–229.
- Mitrenina O.* Correlatives: Evidence from Russian // *Proceedings of FDSL 7.5*. — Peter Lang, 2010. — Pp. 135–151.
- Pietraszko J.* The correlative configuration in Polish // *Slavic Grammar from a Formal Perspective. The 10th Anniversary FDSL Conference, Leipzig 2013*. — Peter Lang, 2015. — Pp. 367–382.
- Tsedryk E.* (Cor)relativization as a by-product of *wh*-probing: The case of Russian. — URL: <https://www.academia.edu/34402600>.
- Vries M. de* The syntax of relativization: PhD thesis / de Vries Mark. — University of Amsterdam, 2002.
- Yanovich I.* Choice-functional series of indefinite pronouns and Hamblin semantics // *Proceedings of SALT*. Vol. 15. — 2005. — Pp. 309–326.
- Богданова-Бегларян Н. В.* Агрессивный узус — или эволюция языковой нормы? // *Верхневолжский филологический вестник*. — 2015. — № 1. — С. 25–30.
- Лютикова Е. А.* Загадки русских относительных предложений. — Апрель 2008. — Материалы к докладу на конференции «Синтаксические структуры — 2».
- Митренина О. В.* Синтаксис коррелятивных конструкций русского языка с позиций генеративной грамматики // *Труды международной конференции «Диалог 2008»*. — 2008. — С. 356–360.
- Холодилова М. А.* Стратегия релятивизации с союзным словом *кто*. — 2010. — Неопубликованный хендаут доклада.