

Head last to head first and left peripheries: evidence from Khanty and Udmurt relatives

Éva Dékány¹ and Orsolya Tánzos^{1,2}

dekany.eva@nytud.mta.hu, orsolyatan@gmail.com

¹Research Institute for Linguistics, HAS, ²Péter Pázmány Catholic University

ICHL 22, 27-31 July 2015

1. Introduction

1.1 OV/VO and RelN/NRel

The OV/VO parameter is a predictor of other word order correlations (table from Croft 2003: 72; see also Greeberg 1963, Lehman 1973, Vennemann 1974, Hawkins 1983, Dryer 1992):

	OV	VO		OV	VO
Clausal orders	SV	VS	Phrasal orders	Post	Prep
	Vaux	AuxV		GN	NG
	VAdv	AdvV		RelN	NRel
	VSubr	SubrV		AN	NA
	PurpV	Vpurp		DemN	NDem
	OcompV	VOcomp		NumN	NNum
	SentQ	Qsent		AdvA	AAAdv

Khanty and Udmurt: SOV languages currently undergoing a shift to SVO. In the wake of the OV to VO shift, other word order parameters are also undergoing a change. Of these, we focus on the RelN/NRel parameter.

RCs before the change:	prenominal	gap strategy	non-finite
	↓	↓	↓
RCs after the change:	postnominal	overt relativizer	finite

1 Our names are in alphabetical order. This material is based upon work supported by the Hungarian Scientific Research Fund under grant OTKA 112057 (Hungarian Generative Diachronic Syntax 2). Dékány's work was also supported by a postdoctoral grant of the Hungarian Academy of Sciences. We thank Márta Csepregi, Katalin Gugán and Mária Sipos for discussion of the Khanty data, Diana Vakhrusheva for Udmurt data, and Alina Duboveckaja for clarifying what is an (im)possible relative clause in Russian.

The changes in the 3 parameters take place in a specific order → it is true for both languages that some logically possible combinations are not attested.

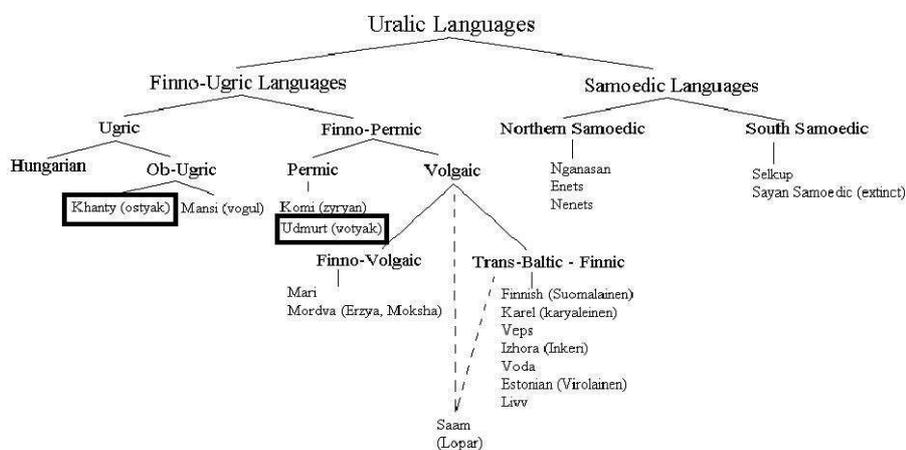
Aim: syntactic analysis of the change at each step, explanation of the missing patterns

Claim: changes are driven by the formation of a left periphery in the RC

1.2 About Khanty and Udmurt

Similarities:

- Uralic, Finno-Ugric languages²
- agglutinative SOV
- one finite verb per sentence, widespread use of non-finite subordination



- RCs are non-finite, prenominal, and use the gap-strategy
- minority languages in the Russian Federation → intensive influence of Russian

Differences:

- in different branches of the Uralic Family: Udmurt is Permic, Khanty is Ob-Ugric
- areal differences: Khanty is spoken in Western Syberia, along the river Ob and its tributaries, Udmurt is spoken in the Volga-Kama Region, just south of the Ural mountains
- different contact languages in addition to Russian: Tatar (SOV) for Udmurt, Nenets and Komi-Zyrian (both SOV) for Khanty
- Udmurt is vulnerable, Khanty is severely endangered (based on the Russian census in 2010, 30 943 total Khanty ethnic population, of which 9600 native speakers; 552 299 total Udmurt ethnic population, of which 339 800 native speakers)³

² language tree from

<http://www.policy.hu/filtchenko/Documenting%20Eastern%20Khanty/Eastern%20Khanty%20Map.htm>

3. RCs in Russian

Russian has 3 types of RCs.

A) prenominal non-finites (Hendery 2012: 202)

- (3) Saša otpravil [napisannoe včera] pis'mo.
Sasha .NOM PRT.send.PST.3SG PRT.write.PASS.PST.PTC.NEU yesterday letter.ACC
'Sasha sent the letter written yesterday.'

B) more commonly postnominal non-finites (Hendery 2012: 202)

- (4) Saša otpravil pis'mo [napisannoe včera].
Sasha .NOM PRT.send.PST.3SG letter.ACC PRT.write.PASS.PST.PTC.NEU yesterday
'Sasha sent the letter written yesterday.'

C) most typically postnominal, finite RCs with a *wh*-based relative operator in spec, CP.

The relativizer agrees with the antecedent in gender and number but takes case from the relative clause (Bailyn 2012). No P-stranding (Miller and Weinert 1998: 351).

- (5) to, [čego ja bojus'] (6) Čego vy boites'?'
that.NEUT.SG which-GEN I fear what-GEN you.NOM fear
'that which I fear' (Bailyn 2012:116) 'what are you afraid of?'

- (7) pričiny, [po kotorym žeščiny brosjut mužčiny]
reasons by which-DAT.PL women throw men
'reasons for which women leave men' (Bailyn 2012:116)

4. Change in position only

If the change affects the position of RCs (RelN → NRel), but not the finiteness of the relativizing strategy:

-- rejected by Udmurt speakers

-- „highly infrequent” and is „eventually self-repaired into” a prenominal non-finite RC in Khanty (Filchenko 2007: 468)

(11) **qol-pa** mən-l-ən?
 where-ILL go-PRS-2SG
 ‘Where are you going?’ (Filchenko 2010: 385) Khanty

(12) So korkan ik ul-i, [**kytyn** lu-ono mynym]
 3SG house.INESS same live-PST.3SG where be-PRT 1SG.DAT
 ‘He lived in the same house, where I have to live.’ Udmurt

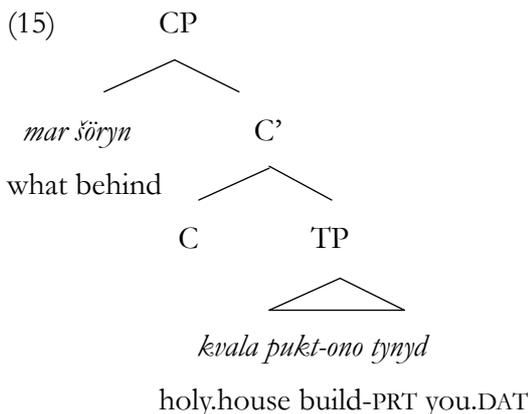
(13) **Kytyn** so ul-i?
 where 3SG live-PST.3SG
 ‘Where did (s)he live?’ Udmurt

Evidence that these are operators, not relative C heads at this stage: may be modified by prepositions, can take plural and case marking (see section 7 for Khanty examples).

(14) So korkan ul-i, [**mar šör̄yn** kvala pukt-ono tynyd]
 3SG house.INESS live-PST.3SG what behind holy.house build-PRT you.DAT
 ‘He lived in the house behind which you have to build the holy house.’ Udmurt

Proposal:

- the *wh*-based relative operator appears here because post-nominal RCs have a left periphery, i.e they project a CP layer
- when the CP layer is present, there is a need to overtly mark clause-typing
- Khanty and Udmurt have no relative complementizers
 → clause-typing is taken care of by a relative operator in spec, CP



6. Excursus on the possibility of mirror-translation

Our approach: change internal to Udmurt and Khanty grammar

Possible position: no real change in the Udmurt/Khanty clause structure, these are Russian structures with Udmurt/Khanty words (i.e. reverse of the typical relexification scenario, see Bakker 2000 on Sri Lanka Malay).

However,

1) Russian has no postnominal non-finites with relative operators; Udmurt and Khanty do

- (16) *Saša otpravil pis'mo, [kotoroe včera napisannoe].
Sasha.NOM PRT.send.PST.3SG letter.ACC which.ACC yesterday PRT.write.PASS.PST.PTC.NEU
'Sasha sent the letter written yesterday.'

→ can't be translations

2) Russian has postnominal non-finite RCs without a relativizer; Udmurt and Khanty don't

- (17) Saša polučil posylku, [otpravlennuju včera].
Sasha.nom get.pst.3sg package.acc.fem prt.send.pass.pst.ptc.acc.fem yesterday
'Sasha got the package that was sent yesterday.'

→ why are these structures not mirror-translated?

3) Russian non-finite RCs have number and case agreement; the Khanty and Udmurt RCs of this type don't (Csepregi 2012: 86)

- (18) Vybraem [samuju čitaemuju] knigu goda
choose.pst.1pl most.acc.fem read.pass.pst.ptc.acc.fem book.acc.fem year.gen
'We chose the most read book of the year.'

→ why are these structures not mirror-translated?

4) Khanty develops relative operators from demonstratives, too, which cannot be explained this way

→ can't be translations

7. The Dem-based relativizer in Khanty

7.1 The system of demonstratives

demonstratives in Khanty:

- proximal vs distal, adnominal vs pronominal, definite/visible vs indefinite/invisible
- adnominally uninflected, inflected in the anaphoric and deictic use

The system of Khanty demonstratives (Surgut dialect, Márta Csepregi, pc.):

	specific		abstract	
	pronominal	adnominal	pronominal	adnominal
proximal	<i>tēmi</i>	<i>tēm</i>	<i>t'it</i>	<i>t'i</i>
distal	<i>tomi</i>	<i>tom</i>	<i>t'ūt</i>	<i>t'ū</i>

Normally, *t'ū* appears adnominally and *t'ūt* pronominally (23),

- (19) a. *mā t'ū rȳt-nat mən-Λ-əm*
 I that boat-COM go-PRS-2SG
 'I take that boat.' (Márta Csepregi, pc.)

- b. *mā t'ūt-nat mən-Λ-əm*
 I that-COM go-PRS-2SG
 'I take that.' (Márta Csepregi, pc.)

However, *t'ū* also occurs pronominally as a complement to Ps (24a) and in object position (24b), but not in oblique case positions → it is uninflected in all its pronominal uses, too

(20)

- a. ***t'ū pyrñə*** *tam miša-nə wās-kən noq náat'-Λ-i-γən.*
 that after there Misa-LOC duck-DU PRT pluck-PRS-PASS-DU3
 'After that the two ducks are plucked by Misa.' (Csepregi 1998: 60)

- b. *nüŋ tom torəm tom iΛəm pälək-nə kăw,*
 you DET world DET front half-LOC stone

- c. ***t'ū küč pow-Λ-e,*** *muγti wīčəpə əntə pitə-Λ.*
 DEM while blow-PRS-SG2.OBJSG through forever not get-PRS(SG3)
 'At the front part of the world [there is] a stone, even though you blow it, you never get through it.' (Csepregi 1998: 64)

7.2 The demonstrative-based relativizer

The distal demonstrative *tʃu/tu/t'ũ* grammaticalized into an element introducing the RC

- (21) *pirəš iki, [t'u lʷw əwi-ʌ-at ma nəmʌʌɣt-əɣəʌ-t-am]*
 old man that 3SG daughter-3SG-INSF 1SG think-FREQ-PRT.PRS-1SG
 'the old man whose daughter I am thinking about' (Csepregi 2012: 87)

Relative operator or relative C?

- inflection doesn't help, *tʃu/tu/t'ũ* is uninflected as a pronoun as well as a relative element (see Potanina 2013: 79 for the latter claim)

→ compatible with both operator and head status

- no reported co-occurrence of wh-based relativizer and dem-based relativizer

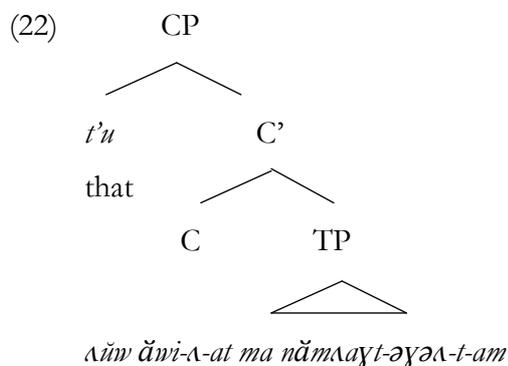
→ could be due to the Doubly Filled Comp filter

- we take the dem-based relativizer to be an operator because it is a recent addition to the language

→ the first step of the relative cycle involves an operator, the cycle probably has not had time to proceed to the operator to C head reanalysis step

Proposal:

- the *tʃu/tu/t'ũ* at the beginning of the RC is a relative operator
- the *wh*-based relativizer and the dem-based relativizer are two competing strategies to mark clause-typing in the relative clause



3SG daughter-3SG-INSF 1SG think-FREQ-PRT.PRS-1SG

Hypothesis based on what happened in Old English and Old Hungarian: reanalysis started in contexts where the pronominal use of that *tʃu/tu/t'ü* was immediately followed by the RC.

The following factors could play a role in the reanalysis:

- 1) always uninflected
- 2) object drop

8. Change in finiteness in post-nominal RCs

Khanty: the relativizer in finite clauses is near-obligatory (Csepregi 2012: 87).

- (23) merəm-qən [muyulə-γən jateswe-wəl aŋk-im]
 tale-DU which-DU tell-PRES.3SG mother-POSS.1SG
 ‘the tales that are told by my mother.’ (Filchenko 2010: 302) present, *wh*-rel
- (24) puγəλ, [mətapi-nə ma səm-a pit-əm]
 village which-LOC 1SG eye-LAT fall-[PST]1SG
 ‘the village where I was born’ (Csepregi 2012: 88) past, *wh*-rel
- (25) mä amə-γal-əm qat [tʃu qarəŋ-nə aməs-wəl]
 1sg sit-PST-1SG house dem bank-LOC sit-PRS.3SG
 ‘I built the house which is on the riverbank.’ (Potanina 2013: 79) present, dem-rel

Udmurt: relativizer is obligatory

- (26) veras’ki todmo-nenym [kudiz jarat-e/jarat-i koçysh-jos-ty]
 talk-PST.1SG friend-POSS.1SG.INS REL.NOM like-PRS.3SG/ like-PST.3SG cat-PL-ACC
 ‘I talked to my friend who likes/liked cats.’

Proposal:

- finite clauses always have a left periphery, they can’t be as truncated as non-finites
- Khanty strongly prefers, while Udmurt requires marking of clause typing
- an overt element on the left periphery is strongly preferred in Khanty and obligatory in Udmurt

→ in absence of a relative complementizer, the relative operator must be used

Khanty RCs without a relativizer (note that the DEM element is not the *tʃu/tu* DEM):

- (27) *män-nə onəl-l-əṃ, tom qu jo-wəl*
1SG-LOC know-PRS-1SG det man walk-PRS.3SG
'I know the man, who is walking there.' (Filchenko 2010: 500)⁵

Why is there no operator here?

Possibility No1: finite inflection may be enough to mark clause size in Khanty

Possibility No2: Filchenko (2010) claims that these are internally headed relative clauses, but provides no evidence

9. The left periphery of Udmurt RCs with a relativizer

Rizzi's (1997, 1999) split CP:

- (28) FORCE (TOP*) INT (TOP*) FOC (TOP*) FIN IP

- Co-occurrence with complementizers
-- clause-final Fin complementizer *shuisa* 'that':

- (29) *Mon todisko co pinalez, [kudze Sasha uramish adziz (*shuisa)].*
1SG know.PRS that child.ACC which.ACC Sasha.NOM street.ABL see.PST.3SG that
'I know that child which Sasha saw on the street.' Udmurt

-- clause-initial Force complementizer *sto* 'that':

- (30) *Mon todisko co pinalez, [(*)sto kudze (*)sto Sasha uramish adziz].*
1SG know.PRS that child.ACC that which.ACC that Sasha.NOM street.ABL see.PST.3SG
'I know that child which Sasha saw on the street.' Udmurt

→ no conclusion can be drawn

⁵ The subject of (31) is „that man who” (Márta Csepregi, pc).

- Topicalization above the operator:⁶

-- subject topic

- (31) Mon todisko so pinalez, [**(*Sasha) kudze** Sasha uramish adziz].
 1SG know.PRS that child.ACC Sasha.NOM which.ACC Sasha.NOM street.ABL see.PST.3SG
 ‘I know the child that Sasha saw on the street.’ Udmurt

-- object topic

- (32) *Mon todisko so pinalez, [**Sashajez kudiz** uramish adziz.]
 1SG know.PRS that child.ACC Sasha.ACC which.NOM street.ABL see.PST.3SG
 ‘I know the child that Sasha saw on the street.’ Udmurt

-- adjunct topic

- (33) *Mon todisko so pinalez, [**Izhkaryn kudiz** Sashajez uramish
 1SG know.PRS that child.ACC Izhevsk.INESS which.NOM Sasha.ACC street.ABL
 adziz.]
 see.PST.3SG
 ‘I know the child that Sasha saw on the street in Izhevsk.’ Udmurt

→ no conclusion can be drawn

- FOC: not applicable due to the characteristics of Focus in Udmurt
- INT: not applicable

Proposal: the relativizer in the spec of the lower CP (FinP); for speakers who don't accept topicalization the higher phrases in the left periphery are possibly not projected or CP is not split

6 Compare Russian: topicalization is not possible, the general complementizer cannot appear in RCs. Oleg Belyaev (p.c.) informs us that topicalization is possible in colloquial Russian, however.

- (i) a. *Eto tot dom, [Saša kotoryj v prošlom godu postroil].
 this that house.NOM Sasha.NOM which.ACC in last.PREP year.PREP build.PST.3SG
 b. * Eto tot dom, [što kotoryj Saša v prošlom godu
 this that house.NOM that which.ACC Sasha.NOM in last.PREP year.PREP
 postroil].
 build.PST.3SG
 ‘This is the house that Sasha built last year.’

10. Conclusions

Original structures: prenominal, non-finite, gap strategy

Steps of the change:

1. prenominal to postnominal
2. gap to relativizer strategy
3. non-finite to finite

Unattested combinations:

- A. prenominal and finite and/or has relativizer
- B. postnominal finite without relativizer in Udmurt
- C. postnominal nonfinite without a relativizer in Udmurt

Main claims:

- I. postnominal RCs in Khanty and Udmurt developed a left periphery
- II. Khanty strongly prefers to overtly mark the left periphery for clause typing
- III. Udmurt makes this marking obligatory
- IV. marking of clause typing is done via relative operators in spec, CP
- V. these operators are grammaticalizing from *wh*-elements (in both languages) and a demonstrative (in Khanty)

Appendix. Diachronic depth of the new RCs and intra-speaker variation

A1. Diachronic depth: When did these structures emerge?

Khanty: reported from the 1950's-1960's by Gulya in headless relative clauses:

(33) [möyöli mänä mas-wəl] t'u məji-γilə-γas
what 1SG-LOC need-3SG that give-TR-PST3.3SG
'What I need, that he gave me.' (Gulya 1966: 86)

(34) tōγ-l-a, [qo wəl-ət tʃu jaγ]
DET-3SG-ILLAT where live-PST.3PL det people
'there, where those people lived (Filchenko 2007, citing Kalinina 1970)

As reported by Csepregi (1983), Karjalainen (1964) contains one sentence with a *wh*-based relativizer; Kalinina (1966, 1970) also contain a few examples. However, all but one of Kalinina's examples are translations from Russian, where the original Russian sentence also contains such a relativizer.

Csepregi (1983): the Kalinina texts have 8 RCs introduced by *when*, 7 of which are non-finite and 1 is finite

Filchenko (2010: 508): 80% participial predicates in RCs, 20% finite predicates

Filchenko (2010: 499): 15% of RCs is introduced by *wh*-based relativizers

Udmurt: relativizers reported by Winkler (2001). *Kud* 'which' was the first *wh*-word to grammaticalize: today it's restricted as a *wh*-word; more characteristically used as relative pronoun,⁷ but still an operator (takes number and case marking and can be modified by Ps).

(35) [Kud-jos-ez-lâ pin'al'-l'os-lâ mon vož-me pot-i], soos pegž'-izâ.
which-PL-DET-DAT child-PL-DAT 1SG anger-1SG.ACC come.out-PST.3SG 3PL run-PST.3.PL
'The boys, which I got angry at, have run away.' (Belyaev 2012, ex. 14)

(36) Mon so korkain uly, [kudiz söryn tyala aryn kvala
1SG that house.INESSlive.PST.1SG which behind next year.INESS holy.house
puktozy].
build.FUT.3SG
'I lived in the house behind which they will build a holy house.'

⁷ In this function *kud* is compounded with a demonstrative suffix *-iz* (see also Suihkonen 2005).

A.2 Intra-speaker variation

3 steps of the change: 1) change in position, 2) change in relativizing strategy (introduction of an operator), 3) change in finiteness.

The original structures are still highly preferred by self-conscious language users and 'purists'. The three varieties live side by side, the same speaker may produce all three variants

→ no separation of the varieties in time or by dialect/idiolect

References

- Alatyrev V. I. 1974 *Грамматика современного удмуртского языка. Синтаксис сложного предложения*. [The grammar of the modern Udmurt language. Syntax of the complex sentences] Ижевск: Удмуртия.
- Bailyn, J. F. 2012. *The syntax of Russian*. Cambridge: CUP.
- Bakker, P. 2000. Rapid language change: creolization, intertwining, convergence. In: Colin Renfrew, A. MdMahon and L. Trask (eds.), *Time depth in historical linguistics*, Vol 2., 585–620. Cambridge: McDonald Institute for Archeological Research.
- Belyaev, O. 2012. Коррелятивная конструкция и конструкция относительного предложения с внутренней вершиной в бесермянском диалекте удмуртского языка [Correlatives and internally-headed relatives in Beserman Udmurt] In: Кузнецова А.И., Сердобольская Н.В., Толдова С.Ю., Сай С.С., Калинина Е.Ю. (eds.). *Финно-угорские языки: фрагменты грамматического описания. Формальный и функциональный подходы*. 647–680.
- Croft, W. 2003. *Typology and universals*. 2nd ed. Cambridge: CUP.
- Csepregi, M. 1983. Rendszerkényszer. Újabb keleti osztják nyelvi jelenségek. In: G. Bereczki and P. Domokos (eds.): *Uráliszttikai tanulmányok Hajdú Péter 60. születésnapja tiszteletére*, 73–84. Budapest: ELTE.
- Csepregi, M. 1996. Orosz nyelvi hatás a déli osztják szintaxisban. In: Mészáros Edit (ed.), *Ünnep könyv Mikola Tibor tiszteletére*. Szeged: JATE Finnugor Tanszék.
- Csepregi, M. 1998. Szurguti osztják chrestomathia. Szeged: JATE.
- Csepregi, M. 2012. Participiális jelzős szerkezetek két hanti nyelvjárásban. *Nyelvtudományi Közlemények* 108: 61–94.
- Dryer, M. S. 1992. The Greenbergian Word Order Correlations. *Language* 68: 81–138.
- Filchenko, A. Y. 2007. *A grammar of Eastern Khanty*. PhD dissertation, Rice University.
- Filchenko, A. Y. 2012. Continuity of information structuring strategies in eastern Khanty: definiteness/topicality. In: P. Suihkonen, B. Cornie and V. Solovyev (eds.), *Argument Structure and Grammatical Relations*, 115–132. Amsterdam and Philadelphia: John Benjamins.
- Gelderen, E. van. 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.
- Gelderen, E. van. 2009. Renewal in the Left Periphery: Economy and the Complementiser Layer. *Transactions of the Philological Society* 107(2): 131–195.
- Greenberg, J. H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: J. H. Greenberg, (ed.) *Universals of grammar*, 73–113. Cambridge, Mass.: MIT Press.

- GSUJa = *Грамматика современного удмуртского языка. Фонетика и морфология*. [The grammar of the modern Udmurt language]. 1962. Перевошиков П. Н., Вахрушев В. М., Алатырев В. И. (eds.), Ижевск: Удмуртское книжн. издво
- Hawkins, J. A. 1983. *Word Order Universals*. New York: Academic Press.
- Heine, B. and T. Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP.
- Hendery, R. 2012. *Relative clauses in time and space*. Amsterdam and Philadelphia: John Benjamins.
- Номлјак 2002. А. Р. Хомляк, Арем-моньцэм ел ки манл... Если моя сказка-песня дальше идёт... Групп Полиграфист, Ханты-Мансийск
- Hopper, M. and E. Traugott (1993) *Grammaticalization*. Cambridge: CUP.
- Kalinina I. M. 1966. Hantyskaja skazka. *Jazyki i toponimija Sibiri* I: 88–95.
- Kalinina I.M. 1970. Kolym pay "tri syna" (hantyskaja skazka). *Jazyki i toponimija Sibiri* III: 45–60.
- Lehmann, W. P. 1973. A Structural Principle of Language and Its Implications. *Language* 49(1): 47–66.
- Nikolaeva, I. 1999. *Ostyak*. Lincom Europa: Muenchen.
- Potanina, O. 2008. Grammaticalization of relative clauses in Eastern Khanty. In *Subordination and coordination strategies in North Asian languages*, 77-84. Amsterdam and Philadelphia: John Benjamins.
- Potanina, O. 2013. Manifestations of language attrition among Eastern Khanty speakers. *Tomsk Journal of Linguistics and Anthropology* 1(2): 75–82.
- Rizzi, L. (1997) The Fine Structure of the Left Periphery. In: Liliane Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.
- Rizzi, L. 2001. On the position Int(errogative) in the left periphery of the clause. In: G. Cinque & G. Salvi (eds.), *Current studies in Italian syntax. Essays offered to Lorenzo Renzi*, 287–296. Oxord: Elsevier.
- Suihkonen, P. 2005. On the Categories and Functions Developed from the Possessive and Deictic Suffixes in Udmurt In: C. Hasselblatt, E. Koponen, A. Widmer (eds.). *Libkkun Lebkos! Beitrage zur Finnougristik aus Anlas des sechzigsten Geburtstages von Hans-Hermann Bartens*, 401–427. Wiesbaden: Harrassowitz.
- Vennemann, T. 1974. Topics Subjects, and Word Order: From SXV to SVX via TVX. In: J. M. Anderson, & C. Jones (eds.) *Historical Linguistics I*, 339–376. Amsterdam, Oxford: North-Holland.
- Winkler, E. 2001. *Udmurt*. Muenchen: Lincom.