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):

<table>
<thead>
<tr>
<th>OV</th>
<th>VO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clausal orders</td>
<td>OV</td>
</tr>
<tr>
<td>SV</td>
<td>VS</td>
</tr>
<tr>
<td>Vaux</td>
<td>AuxV</td>
</tr>
<tr>
<td>VAdv</td>
<td>AdvV</td>
</tr>
<tr>
<td>VSubr</td>
<td>SubrV</td>
</tr>
<tr>
<td>PurpV</td>
<td>Vpurp</td>
</tr>
<tr>
<td>OcompV</td>
<td>VOcomp</td>
</tr>
<tr>
<td>SentQ</td>
<td>Qsent</td>
</tr>
</tbody>
</table>

| Phrasal orders | OV | VO |
| Post | Prep |
| GN | NG |
| RelN | NRel |
| AN | NA |
| DemN | NDem |
| NumN | NNum |
| AdvA | AAdv |

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 languages
- agglutinative SOV
- one finite verb per sentence, widespread use of non-finite subordination

**Differences:**
- in different branches of the Uralic Family: Udmurt is Permic, Khanty is Ob-Ugric
- areal differences: Khanty is spoken in Western Siberia, 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)

---

• Khanty has morphologically unmarked objects, Udmurt has differential object marking (only
definite direct objects bear Acc, other objects are unmarked)

• Khanty is a dialect continuum with 3 main dialects: western (a.k.a northern), eastern, and the
extinct southern

Current situation:

• diglossia, unidirectional bilingualism (cca. 100%)

• Khanty: most living speakers went to boarding school, which helped assimilation; young people
  speak primarily Russian, children learn Khanty only if their parents have very traditional jobs
  (Sipos 2014)

• Udmurt: many children learn the language but only those living in remote rural areas continue to
  use it actively (Ethnologue)

• strong influence of the inflectional SVO Russian on both the lexicon and the syntax

• spread of OV and finite subordination, but the two languages are at different stages of the
  OV – VO change

2. Original RCs in Khanty and Udmurt

Prenominal, non-finite RCs employing the gap strategy (Khanty: Nikolaeva 1999, Filchenko 2007, 2012,
Prenominal RCs cannot be finite and cannot have a relative operator.

(1)  [katüka-m-am]  kuŋ  put-nů  kit’
    catch-PTC.PST-1SG fish pot-LOC stay-[PST.3SG]
    ‘The fish that I have caught stayed in the pot.’ (Csepregi 2012, ex. 9b)    Khanty

(2)  Sasha  [pes’atajen  puktem]  korkan  kyk  ar  ule  in’i
    Sasha.NOM grandfather.INSR built.PRT house.INESS two year live.PRS.3SG already
    ‘Sasha has been living in the house that was built by his grandfather for two years.’    Udmurt

Claim: these RCs are no bigger than TP → relative operators sit in spec, CP; that position is not
projected in these RCs.

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 PST.NEUT.3SG.pst 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 PST.NEUT.3SG.pst 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,      [čeго 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 brosajut 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)
Proposition: post-nominal RCs in these languages have a left periphery, which needs overt marking.

5. Change in position as well as relativizing strategy

Post-nominal non-finite RCs are possible if they feature a relativizer.

5.1 The relative cycle

Relativizers often grammaticalize from *wh*-pronouns or demonstrative pronouns (Van Gelderen 2004, 2009, Hopper and Traugott 1993, Heine and Kuteva 2002), and the relative operator may later grammaticalize into a C head (and further grammaticalize into a higher C head), a process known as the Relative Cycle (Van Gelderen 2004, 2009).

\[ \text{interrogative pronoun} \rightarrow \text{relative pronoun} \rightarrow \text{relative complementizer} \rightarrow \text{higher C head} \]

\[ \text{demonstrative pronoun} \]

5.2 The *wh*-based relativizer

In post-nominal RCs both Khanty and Udmurt may feature a *wh*-based relative operator. 4

(10) ju wül-wül qa-nü [qo mā wül-m-ām]
3SG live-PRES.3SG house-LOC where 1SG live-PST.PRT-1SG

‘He lives in the house where I lived.’ (Potanina 2013: 79)

Khanty

---

4 The *wh*-based relative operator can also be found in the Northern Khanty dialect:

(i) nin, lop-əʌ, xọlta man-ə-am? DU2 say-PRES.SG3 where.to go-PRES-DU2

‘He says, where are you going?’ (Homlijak 2002)

(ii) min man-ə-amən, lop-əʌ, [xọlta pa jöxt-ə-amn], s’iw DU1 go-PRES-DU1 say- PRES.SG3 where.to PTCL arrive-PRES-DU1 there man-ə-amən.
go-PRES-DU1

‘We are walking, he says, where we arrive, there we go!’ (Homlijak 2002)
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).

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

\[(16) \text{ *Saša} \text{ otpravil} \text{ pis’mo, [kotoroe} \text{ včera napisannoe].} \]
\[
\begin{align*}
\text{Sasha.NOM} & \text{ PRT.send.PST.3SG letter.ACC which.ACC yesterday PRT.write.PASS.PST.PTC.NEU} \\
\text{‘Sasha sent the letter written yesterday.’}
\end{align*}
\]

→ can’t be translations

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

\[(17) \text{ Saša} \text{ polučil} \text{ posylku, [otpravlennuju} \text{ včera}.] \]
\[
\begin{align*}
\text{Sasha.nom} & \text{ get.pst.3sg package.acc.fem prt.send.pass.pst.ptc.acc.fem yesterday} \\
\text{‘Sasha got the package that was sent yesterday.’}
\end{align*}
\]

→ 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)

\[(18) \text{ Vybraem [samuju čitaemuju] knigu goda} \]
\[
\begin{align*}
\text{choose.pst.1pl most.acc.fem read.pass.pst.ptc.acc.fem book.acc.fem year.gen} \\
\text{‘We chose the most read book of the year.’}
\end{align*}
\]

→ 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, adnominally vs pronominally, definite/visible vs indefinite/invisible
- adnominally uninflected, inflected in the anaphoric and deictic use

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

<table>
<thead>
<tr>
<th>specific</th>
<th>abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pronominal</td>
</tr>
<tr>
<td>proximal</td>
<td>tēmi</td>
</tr>
<tr>
<td>distal</td>
<td>tom</td>
</tr>
</tbody>
</table>

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

(19)  a. mā t'ū rŷt-nat mən-a-əm
    I that boat-COM go-PRT-2SG
    ‘I take that boat.’ (Márta Csepregi, pc.)

    b. mā t'ūt-nat mən-a-əm
    I that-COM go-PRT-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 pyrna tam miša-nə wāś-kən noq ňāt'-A-i-γən.
    that after there Misa-LOC duck-DU PRT pluck-PRT PASS-DU 3
    ‘After that the two ducks are plucked by Misa.’ (Csepregi 1998: 60)

    b. nüng tom torəm tom iaəm pālək-ə kəw,
    you DET world DET front half-LOC stone

    c. t'ū kūč pow-ə-e, muñti wĩčəpə əntə pitə-ə.
    DEM while blow-PRT-SG2.OBJSG through forever not get-PRT(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 ɟʧu/tu/t'ū grammaticalized into an element introducing the RC

(21) pirəš iki, [t'u āw āwi-ʌ-at ma nāmagt-ɵma-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, ɟʧu/tu/t'ū 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 ɟʧu/tu/t'ū 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

(22) CP
    t'u C'
    that
    C TP
    āw āwi-ʌ-at ma nāmagt-ɵma-t-am
    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* *ʧʰu/tʰu/tʰu* 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

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

(23) merən-qan [muŋu-la-γən jateswe-wal 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) puŋə, [maŋapi-na ma sām-a pit-am]
village which-LOC 1SG eye-LAT fall-[PST]1SG
‘the village where I was born’ (Csepregi 2012: 88) past, *wʰ*-rel

(25) mā amə-yal-am qat [ʧu qajən-na 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 kochysh-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
- Khanyt strongly prefers, while Udmurt requires marking of clause typing
- an overt element on the left periphery is strongly preferred in Khanyt 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 $g'$u/$tu DEM):

(27) män-ə onəl-əm, tom qu jo-wał

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.’

  -- 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.’

→ 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, [(\text{*Sasha}) kudze Sasha uramish adziz].
\(1\text{SG} \text{know:PRS that child.ACC Sasha.NOM which.ACC Sasha.NOM street.ABL see.PST.3SG}\)

‘I know the child that Sasha saw on the street.’

Udmurt

-- object topic

(32) *Mon todisko so pinalez, [Sashajez kudiz uramish adziz.]
\(1\text{SG} \text{know:PRS that child.ACC Sasha.ACC which.NOM street.ABL see.PST.3SG}\)

‘I know the child that Sasha saw on the street.’

Udmurt

-- adjunct topic

(33) *Mon todisko so pinalez, [Izhkaryn kudiz Sashajez uramish adziz.]
\(1\text{SG} \text{know:PRS that child.ACC Izhevsk.INESS which.NOM Sasha.ACC street.ABL see.PST.3SG}\)

‘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.

1. a. *Eto tot dom, [Saša kotoryj v prošlom godu postroil].
\(\text{this that house.NOM Sasha.NOM which.ACC in last.PREP year.PREP build.PST.3SG}\)

b. *Eto tot dom, [što kotoryj Saša v prošlom godu postroil].
\(\text{this that house.NOM that which.ACC Sasha.NOM in last.PREP year.PREP build.PST.3SG}\)

‘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 is 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 wb-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) [mɔyoli mänä mas-wə] t’u məjį-ɣilə-ɣas
what 1SG-LOC need-3SG that give-TR-PST.3SG
‘What I need, that he gave me.’ (Gulya 1966: 86)

(34) tɔɣ-1-a, [qo wəl-ə tʃ u ją]
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-lə pın’al-’ləs-lə mon vəz-me pot-i], soos pęg3’i-izə.
which-PL-DET-DAT child-PL-DAT 1SG anger-1SG.ACC come.out-PST.3SG 3PL run-PST.3PL.
‘The boys, which I got angry at, have run away.’ (Belyaev 2012, ex. 14)

(36) Mon so korkain uly, [kudiz sɔrən tyala aryn kvala
1SG that house.INESS live.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


Грамматика современного удмуртского языка. Фонетика и морфология. [The grammar of the modern Udmurt language]. 1962. Перевощиков П. Н., Вахрушев В. М., Алатырев В. И. (eds.), Ижевск.


Homljak 2002. А. Р. Хомляк, Арём-моньщём ел ки манал... Если моя сказка-песня дальше идёт...

Групп Полиграфист, Ханты-Мансийск


