1. Introduction

1.1 SOV to SVO and pre- to postnominal relatives: the phenomenon to be accounted for

It is well-known that 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>Clausal orders</th>
<th>OV</th>
<th>VO</th>
<th>Phrasal orders</th>
<th>OV</th>
<th>VO</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>VS</td>
<td></td>
<td>Post</td>
<td>Prep</td>
<td></td>
</tr>
<tr>
<td>Vaux</td>
<td>AuxV</td>
<td></td>
<td>GN</td>
<td>NG</td>
<td></td>
</tr>
<tr>
<td>VAdv</td>
<td>AdvV</td>
<td></td>
<td>RelN</td>
<td>NRel</td>
<td></td>
</tr>
<tr>
<td>VSubr</td>
<td>SubrV</td>
<td></td>
<td>AN</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>PurpV</td>
<td>Vpurp</td>
<td></td>
<td>DemN</td>
<td>NDem</td>
<td></td>
</tr>
<tr>
<td>OcompV</td>
<td>VOcomp</td>
<td></td>
<td>NumN</td>
<td>NNum</td>
<td></td>
</tr>
<tr>
<td>SentQ</td>
<td>Qsent</td>
<td></td>
<td>AdvA</td>
<td>AAdv</td>
<td></td>
</tr>
</tbody>
</table>

Khanty and Udmurt are originally SOV languages that are 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.

Khanty and Udmurt RCs

before the change: prenominal gap strategy non-finite

\[ \downarrow \quad \downarrow \quad \downarrow \]

after the change: postnominal overt relativizer finite

The change takes place in 3 steps; the 3 individual changes happen in a specific order → it is true for both languages that some logically possible combinations are not attested.

---

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.

---
Aim: syntactic analysis of the change at each step
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
- 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)

---

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.


1. \([\text{katūa-m-am}]\) kūa put-nū 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. \([\text{Kük nunal zorūš’}]\) zor gült-i-z š’rūs’-ez
   two day fall-ING rain.NOM destroy-PST road-ACC
   The rain that has been falling for 2 days destroyed the road.’ (Belyaev 2012, ex. 6a) Udmurt

Prenominal RCs cannot be finite and cannot have a relative operator.
Claim: this is because they are truncated clauses; they don’t have a left periphery. Relative operators sit in spec, CP; that positon is not projected in these RCs.

3. RCs in the contact language

Russian has 3 types of RCs.

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

(4) Saša _otpravil_ [napisannoе 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)

(5) Saša _otpravil_ pis’mo [napisannoе včera].
    Sasha _NOM_ PRT.send.PST.3SG _letter.ACC_ PRT.write.PASS.PST.PTC.NEU yesterday

    ‘Sasha sent the letter written yesterday.’

C) 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).

(6) to, [čego ja bojus’] (8) 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?’

(9) pričiny, [po kotorym žešćiny brosajut mužčiny]
    reasons by which-DAT.PL women throw men
    ‘reasons for which women leave men’ (Bailyn 2012:116)
4. The change to the head-first structure

The head-final to head-first shift first affects the position of RCs (RelN → NRel), but not the finiteness of the relativizing strategy.

(10) kuλ, [ma-nű katλ-ũm] put-nű kiʾr
    fish 1SG-LOC catch-PTC.PST pot-LOC stay-[PST.3SG]

‘The fish that I have caught stayed in the pot.’ (Csepregi 2012, ex. 9c)       Khanty

This type is „highly infrequent” and is „eventually self-repaired into” a prenominal non-finite RC in Khanty (Filchenko 2007: 468) and does not exist in Udmurt. We suggest that this is because post-nominal RCs in these languages have a left periphery, which needs overt marking.

5. The emergence of the relativizer

In post-nominal RCs a relativizer may appear.

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).

(11) interrogative pronoun

    relative pronoun      →  relative complementizer      →  higher C head

    demonstrative pronoun

5.1 The wh-based relativizer

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

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

(i) nin, lop-α, xoATA man-λ-αtan?
    DU2 say-PRES.SG3 where.to go-PRES-DU2
    ‘He says, where are you going?’ (Homljak 2002)

(ii) min man-λ-aman, lop-α, [xoATA pa jøxt-λ-aman], sʾiw
    DU1 go-PRES-DU1 say- PRES.SG3 where.to PTCL arrive-PRES-DU1 there
    man-λ-aman.
    go-PRES-DU1
    ‘We are walking, he says, where we arrive, there we go!’ (Homljak 2002)
(12) 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)  

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

(14) 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.’

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

Supporting evidence that these are operators, not relative complementizers at this stage: may be modified by prepositions, can take plural and case marking (relevant Khanty examples can be found in section 5).

(16) So korkan ul-i, [mar shōryn kvala pukt-ono tnyd]
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.’

Claim:
- the wh-based relative operator may appear here because post-nominal RCs have a left periphery, i.e. they may 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
In which CP does the relativizer sit in Rizzi’s (1997, 1999) split CP?

(18) FORCE (TOP*) INT (TOP*) FOC (TOP*) FIN IP

No co-occurrence with the complementizer *shuisa ‘that’:

(19) 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

No topicalization above the operator:

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

Udmurt

Proposal: (at least in Udmurt), the relativizer in the spec of the lower CP (FinP), the higher phrases in the left periphery are not projected.

5 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 postroil].
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.’
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

(21) *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.’

2) Russian participial clauses have number and case agreement; the Khanty and Udmurt RCs of this type don’t (Csepregi 2012: 86)

(22) 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.’

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

5.2 The Dem-based relativizer in Khanty

demonstratives in Khanty:

- distinguish bw. proximal and distal, as well as definite/visible and 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></th>
<th>specific</th>
<th>abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pronominal</td>
<td>adnominal</td>
</tr>
<tr>
<td>proximal</td>
<td>témi</td>
<td>tém</td>
</tr>
<tr>
<td>distal</td>
<td>tomi</td>
<td>tom</td>
</tr>
</tbody>
</table>

Normally, t’ū appears adnominally and t’ūt pronominally (23 a and b), however, t’ū also occurs as a complement to Ps (23c) and pronominally in subject or object position (23d) → it is always uninflected
The distal demonstrative $t'ū/tu/t'ǔ$ grammaticalized into an element introducing the RC (it is uninflected as a relative operator, too, see Potanina 2013: 79):

(24) piršt iki, [t'ũ 쑨w āwi-l-at ma nāmāŋy-t-γə-γə-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)

Claims:

- the $t'ũ/tu/t'ǔ$ at the beginning of the RC is a relative operator
- the $w$b-based relativizer and the dem-based relativizer are two competing strategies to mark clause-typing in the relative clause
- we suspect that the fact that $t'ũ/tu/t'ǔ$ is always uninflected played a role in its reanalysis, and the reanalysis started in contexts where the pronominal use of that $t'ũ/tu/t'ǔ$ was immediately followed by the RC

(25) CP

\[ \begin{array}{c}
\text{t'ũ} \\
\text{that}
\end{array} \]
\[ \begin{array}{c}
\text{C'} \\
\text{C} \quad \text{TP}
\end{array} \]

\[ \text{Șuŋ āwi-l-at ma nāmāŋy-t-γə-γə-t-am} \]

3SG daughter-3SG-INSF 1SG think-FREQ-PRT-PRS-1SG
6. Change in finiteness

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

(26) merəm-qən [muŋulə-ɣən jateswe-wəl əŋk-im]
    tale-DU which-DU tell-PRES.3SG mother-POSS.1SG
    ‘the tales that are told by my mother.’ (Filchenko 2010: 302)

(27) puɣə, [matapi-na ma sâm-a pit-əm]
    village which-LOC 1SG eye-LAT fall-[PST]1SG
    ‘the village where I was born’ (Csepregi 2012: 88)

(28) mä wel-s-əm [qo köʃkəl qasi käš-äyi wajə ् lək]
    1SG be-PST2-1SG where hunter man find-PST.3SG animal track
    ‘I have been to that place, where the hunter found the animal track.’ (Filchenko 2007: 500)

(29) mä ama-yal-əm ḡat [tʃu qaŋən-na aməs-wəl]
    1sg sit-PST-1SG house dem bank-LOC sit-PRES.3SG
    ‘I built the house which is on the riverbank.’ (Potanina 2013: 79)

Udmurt: relativizer is obligatory

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

(31) veras’ki todmo-nenym [kudiz jarat-i kochysh-jos-ty]
    talk-PST.1SG friend-POSS.1SG.INS REL.NOM like-PST.3SG cat-PL-ACC
    ‘I talked with my friend who liked cats.’

Claim:

- 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 'u/ tu DEM):

(32) mä wel-käs-im kőřøy [ti ni øvø-wel n’an’]
1SG do-PST-1SG knife dem woman cut-PRS.3SG bread
‘I made the knife which that woman cuts the bread with’ (Potanina 2008: 83, 2013:80)
‘I have made the knife, which a woman cuts the bread with’ (Filchenko 2010: 499)

(33) män-nə onəl-l-əm, tom qu jo-wel
1SG-LOC know-PRS-1SG det man walk-PRS.3SG
‘I know the man, who is walking there.’ (Filchenko 2010: 500)

(34) män-nə illə-ne onəl-l-əm, ti quj-āli ajrī-ne jayentə-wel
1SG-LOC ago-LOC know-PRS-1SG det man-DIM canoe-LOC go-PRS.3SG
‘I have known for long the boy, who is going in the canoe.’ (Filchenko 2010: 500)

In (33) and (34) the subjects are „that man who” and „that boy who” (Márta Csepregi, pc) so these are finite RCs without a relative operator. Note, however, that Filchenko (2010) claims that these are internally headed relative clauses, in which case it is not surprising that there is no relative operator in them.

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

7.1 Diachronic depth

When did these structures emerge?
Khanty: reported from the 1950’s-1960’s by Gulya in headless relative clauses:

(35) [møyöli mänä mas-wel] t’u mæji-yilə-yaș
what 1SG-LOC need-3SG that give-TR-PST3.3SG
‘What I need, that he gave me.’ (Gulya 1966: 86)

(36) tøy-l-a, [qo wəl-ər tʃu jy] tŋəl ɣas
DET-3SG-ILLAT where live-PST.3PL det people
‘there, where those people lived (Filchenko 2007, citing Kalinina 1970)
Csepregi (1983) reports that Karjalainen (1964) also contains one sentence with a *wh*-based relativizer; and that 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 reports that in the Kalinina texts there are 8 RCs introduced by *when*, 7 of which are non-finite and 1 is finite.

Filchenko (2010: 508) on Khanty: 80% participial predicates in RCs, 20% finite predicates
Filchenko (2010: 499) on Khanty: 15% of RCs is introduced by *wh*-based relativizers

Udmurt: Winkler (2001) states that in Udmurt interrogative pronouns such as *kin* ‘who’, *mar* ‘what’, etc. function as relative pronouns. However, he also claims that the usage of *kud* ‘which’ is restricted as a *wh*-word and it is more characteristicly used as relative pronoun. In this function *kud* is compounded with a demonstrative suffix –*iz*. Suihkonen (2005) argues that the demonstrative suffix is used to restrict the function of *kud* pronouns to relative pronouns.

As a relative pronoun *kudiz* ‘which’ can be marked with cases or can be merged with postpositions.

(37) **[Kud-jos-ecz-lə] pin’al’-l’os-lə mon vož-me pot-i], soos pegʒ’-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)

(38) **Mon so korkain uly, [kudiz söryn tyala aryn kvala**

1SG that house.INESSlive.PST.1SG which behind next year.INESS holy.house

build.FUT.3SG
‘I lived in the house behind which they will build a holy house.’

### 7.2 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
8. 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 wh-elements (in both languages) and a demonstrative (in Khanty)

References


GSUJa = Грамматика современного удмуртского языка. Фонетика и морфология. [The grammar of the modern Udmurt language]. 1962. Перевоциков П. Н., Вахрушев В. М., Алатырев В. И. (eds.), Ижевск: Удмуртское книжн. издво


Homljak 2002. Л. Р. Хомляк, Арем-моньщэм ел ки манл... Если моя сказка-песня дальше идёт...

Групп Полиграфист, Ханты-Мансийск


