The locative-marked agent in Yugan Khanty active sentences: A data-driven pilot study[1]

Zsófia Schön

LMU Munich, Institute for Finno-Ugric Studies
LMU Munich, Institute for General Linguistics and Linguistic Typology

Az aktív ragozású igei állítmányok mellett megjelenő jelölt ágens (locAg) a hanti érdekes jelensége, hiszen nominatív nyelvekben az alany következetesen nominatívusszal jelölt. A szurguti nyelvjárásban, főleg a jugáni alnyelvjárásban mégis ettől eltérően, lokatívusszal kódolva is megjelenik. Ennek szociolingvisztikai és nyelvtani okait járja körül ez a korpuszalapú tanulmány. Rámutat arra, hogy a jelenség ugyanúgy független az adatközlők idiolektusától vagy szociológiai státuszától, mint az ige szemantikájától, ragozásától, a mondatok szórendjétől vagy a locAg és a tárgy szófajától. Bár a példák információszerkezeti szempontból is lényegesen sokszínűbbek, mint azt az eddigi szakirodalom bemutatta, mégis ez tűnik az egyetlen olyan szempontnak, amely motiválhatja a locAg megjelenését. A jelölt ágensű aktív mondatokban szembetűnő az alany hangsúlyozásának fontossága, valószínűleg szalienciájának jelölése iránti igény. Ugyanakkor a lokatívuszragos ágens nem csak topikális jellegű lehet, és nem kizárólag két konkuráló topik esetén jelenik meg, hanem topikfolytonosság, topikváltás esetén is, valamint kerülhet fókuszpozícióba is.

Keywords: agent, locative, Khanty, Surgut Khanty, Yugan Khanty, information structure

Kulcsszavak: ágens, lokatívusz, hanti, szurguti hanti, jugáni hanti, információszerkezet

---

[1] This is the written and revised version of my talk "A Yugan Khanty corpus, a textbook and the locative-marked agent" given in the talk series of the Uralic Information Centre, Hungarian Research Centre for Linguistics on 23 February 2022. I am grateful for the fruitful comments of Jeremy Bradley, Márta Csepregi, László Fejes, Katalin Gugán, Nikolett F. Gulyás, Gwen Eva Janda, Szilvia Németh, and discussions with Erika Asztalos, Rogier Blokland, Lena Borise, Andrey Filchenko, Ferenc Havas, Katalin É. Kiss, Gerson Klumpp and Ksenia Shagal. Furthermore, I have to thank my anonymous reviewers for their well-founded remarks. Above all, I am indebted to all Khanty native speakers who shared their knowledge with me and made this research possible. My thoughts and feelings are always with them.

Nyelvtudományi Közlemények 118: 89–135.
DOI: 10.15776/NyK.2022.118.3

1. Introduction

In this paper, I intend to make a first step towards an analysis of the locative-marked agent (locAg) in Yugan Khanty active sentences and justify its existence in this (sub)dialect of Khanty. The idea of this research first arose while I was working on my collection of Yugan Khanty tales. Editing the manuscript of a storybook, going through the changes made by the native language consultant, I realized that the use of locAg in active sentences is quite more significant than assumed before.

Khanty (a Finno-Ugric, more specifically, Ob-Ugric language) itself is subject to large dialectal variation. In the Easternmost dialect, the Vakh-Vasyugan dialect, the phenomenon of the locAg in active sentences is well-established and familiar. The language variant under study here is another Eastern Khanty dialect, Surgut Khanty, with a focus on its literary and Yugan Khanty variants. The term *Surgut Khanty* designates various Khanty subdialects spoken mainly along the rivers Yugan, Pim, Tromagan, Agan, Salym and Demyanka, with around 2800 speakers (Csepregi 2017: 17), as well as the standardized, written form that those subdialects share (Schön 2020: 348). Yugan Khanty is the only southern subdialect of Surgut Khanty and has around 900 speakers (Csepregi 2017: 17). When I refer to Yugan Khanty, by this I mean the local variant as it is spoken in the 21st century. If its literary (Surgut Khanty) form is used, the data is always represented in Cyrillic letters.

Additionally, Surgut Khanty, as well as Yugan Khanty, is a nominative, pro-drop language with a dominant SOV word order and a complete singular-dual-plural paradigm in conjugation as well as in declension. It has two verbal moods (active, passive), two verbal tenses (present, past) and two verbal conjugations (subjective, objective). The subject of the sentence is dominantly in the nominative, and finite verbs always show subject agreement. A finite verb in objective conjugation shows subject-object agreement. There exists a nominal case system with nine cases and a pronominal case system with eleven, where the accusative case exclusively occurs with personal pronouns (Schön – Gugán 2022).

In the literature, the locAg is also called “agent” (Kulonen 1989: 301), “Loc-S/A” (Filchenko 2006: 61), “Loc-marked S/A” (Filchenko 2006: 70), “locative subject” (Sosa 2017: 67), “locative-marked subject” (Klumpp – Skribnik 2022: 1024), or “non-canonical Subject” (Sosa – Virtanen 2018: 253). I decided to use the term “locative-marked Agent” (locAg) to emphasize its connection with active (1), (3) and passive sentences (2), where by *agent* I mean the conscious, willful performer of an

action, marked here by the locative case suffix, appearing as the subject of active verbs in this study. Elsewhere, the locAg appears as an agent of a passive sentence besides the grammatical subject.

In example (1), a neutral Yugan Khanty active sentence is shown with a subject (*mɐ:* 'I') in the nominative, a noun modified with a participle as an object (*qɒ:łtə tɒ:ɣi* 'a place to spend the night') and a verb in present tense, subjective conjugation (*kəntʃłəm* 'I'm searching'). (2) is a passive sentence where the verb (*tʃe:βiji:n* 'the two of them were hidden') shows agreement with the (grammatical) subjects, both in dual[2] (*tɯ:łəsɣən* 'two moons' and *qatłɣən* 'two suns'), and the agent is marked with the locative case (*mɛnʲiłnə* 'by his daughter-in-law'; locAg). (3) is an example of the locAg (*nɯŋnə* '(by) you') in active sentences in Yugan Khanty as the verb (*jɐ:ɣlijɛ* 'you poked it out') shows agreement with the personal pronoun in the locative (*nɯŋnə* '(by) you'). The object in (3) is dropped, but the morphological marker on the verb identifies it (objective conjugation; *jɐ:ɣlijɛ* 'you poked it out').

(1) *mɐ: qɒ:ł-tə tɒ:ɣi kəntʃ-ł-əm*
1SG spend_the_night-PTCP.PRS place search-PRS-1SG

'I'm searching for a place to spend the night.'
(Kayukova – Schön 2019)

(2) *pɐ:nə mɛnʲ-ił-nə tɯ:łəs-ɣən*
and daughter_in_law-SG<3SG-LOC moon-DU

qatł-ɣən uː:ł tʃe:βij-i:n
sun-DU down[3] hide.PST-PASS.3DU

'And the moon and the sun were hidden by his daughter-in-law.'
(Kayukova – Schön 2019)

[2] The function of the dual here is not numeral, but connective (Csepregi 2017: 76).

[3] The glosses given here may diverge from the original glosses in the sources due to different glossing strategies, like glossing word forms morpheme-by-morpheme or not combining preverbs with verbs.

(3) *tʲuː mɐː βiːɬɛ juːɣ ont-ɐ*
that 1SG in_fact tree interior-LAT
jɛrəm-ɐm tɐːtn ***nuŋnə***[4] *βiːɬɛ*
stay_too_long.PTCP.PST-1SG during **2SG.LOC** in_fact
kɵtʃəɣ-nɐt ***jɐːɣlij-ɛ***
knife-COM **poke_out.PST-SG<2SG**

'In fact, while I was staying too long inside the tree, you poked it [= my eye] out with the knife.' (Kayukova – Schön 2019)

Realizing that in my Yugan Khanty corpus the locAg in active sentences appears more often than expected, I began to look for similarities among the examples and for different reasons of its existence, like idiolectal (**3.1.**) or dialectal (**3.2.**) motivation, influence from ritual language (**3.3.**), verbal semantics (**3.4.**), transitivity or conjugation (**3.5.**), correlation with word order (**3.6.**), form of the locAg (**3.7.**) and the direct object (**3.8.**), as well as information structure (**3.9.**). The data is summarized in Table 12, where additionally the question of animacy is examined (**3.10.**). Some highly hypothetical explanations of the origin of the locAg in Yugan Khanty active sentences are provided in section (**4.**). The different possible motivations for the usage of the locAg in Yugan Khanty active sentences are summarized, stating their relevance in section (**5.**). The analysis is completed by a comparison of my findings with other research on the same topic in Khanty like Csepregi (2020, 2021), Filchenko (2006), Sosa (2017), and Sosa – Virtanen (2018) for Surgut and/or Vasyugan Khanty, in section (**5.**).

As the grammaticality of the locAg in active sentences is not doubted, I will not discuss that question in the present paper (see Sosa 2017: 183–184). Furthermore, even if the phenomenon of the locAg in active sentences in Eastern Khanty dialects has been described traditionally in the literature as "ergative construction", "ergativity" or lately as "transitive construction", "semi-transitive construction", "transitivity" (e.g., Balandin 1948; Csepregi 2020, 2021; F. Gulyás 2018; Filchenko 2006, 2011; Gulya 1970, 1982, 1994; Havas 2003, 2011; Honti 1971, 1984; Kulonen 1989, 1991; Ruttkay-Miklián 2003), this question is not within the scope of my research, either.

[4] The locAg, the direct object and the corresponding active verb with their glosses are marked in all examples with emboldening.

2. Corpus and data

The data used for the research presented in this paper comes from three different, partly overlapping corpora, handled here as one corpus:

A) The first part of it is the *Yugan Khanty corpus* of the project *Ob-Ugric database: analysed text corpora and dictionaries for less described Ob-Ugric dialects* (OUDB), available online (https://www.oudb.gwi.uni-muenchen.de/?abfrage=YK_corpus). This searchable online database contains a text collection of 56 tales and legends collected by Zsófia Schön and Lyudmila Kayukova in Western Siberia between 2010 and 2016. The transcription represents spoken language by means of a broad phonemicization using IPA symbols, showing speech disfluency, elision, epenthesis, as well as subdialectal features. All 56 texts are glossed, where the glossing follows the Leipzig Glossing Rules. Most of the texts have an English, German and/or Russian translation (39, 47, 29 texts, respectively). The database is enriched by audio material and a lexicon.

B) The second part of the corpus consists of the material of a bilingual Khanty-German storybook *Märchen und Legenden vom Großen und Kleinen Jugan* [*Tales and Legends from the Bolshoy and Malyy Yugan rivers*] (planned title) to be published by *Kulturstiftung Sibirien* in 2023 with Stephan Dudeck as series editor. Schön (2023) contains altogether 84 tales and legends, of which 56 are from the *Yugan Khanty corpus*, 24 from a further collection of Zsófia Schön and Lyudmila Kayukova (2010–2016) and 4 from the Yugan Khanty collection of Zoltán Nagy. All texts of the storybook represent standardized written language in an extended Cyrillic alphabet. The orthography used was established by Agrafena Pesikova and is corrected by her and Lyudmila Kayukova.

While editing the texts, I realized that my language consultant and corrector has added 13 and crossed out 6 locAg in active sentences. These 19 edited examples are neither part of the corpus used in this study nor within the scope of this research and will be examined elsewhere. For the research presented in this paper, only those examples are used where the locAg in active sentences (a) can be heard on the sound recording from the speaker him/herself and at the same time (b) remained in the storybook after editing.

Furthermore, it has to be pointed out that the same 56 texts of the *Yugan Khanty corpus* and the storybook do not only differ in transcription methods, but partly in content, too, in the sense that the version in the *Yugan Khanty corpus* contains additional sentences showing the cir-

cumstances of elicitation or including comments from the speakers as well as dialogues. Nevertheless, there is only one example of the locAg in an active sentence occurring exclusively in the *Yugan Khanty corpus* (4). The 24 examples that occur in both the *Yugan Khanty corpus* and in the storybook are only counted once and cited from the *Yugan Khanty corpus* as (5), (6), (7). The 14 examples which only occur in the storybook are cited in this paper with Cyrillic letters as (10), (13), (18).

C) The third part of the corpus covers two personal accounts excluded from the *Yugan Khanty corpus* and the storybook as well because of their genre. They are part of my *Yugan Khanty Personal Archive*. The two texts were collected in 2015 by Zsófia Schön and Lyudmila Kayukova at the upper stream of the Bolshoy Yugan river (Schön 2015a, Schön 2015b) and will be represented like the texts of the storybook with Cyrillic letters: (9), (12).

To sum up, the corpus used for this research contains 86 narratives (tales, legends and personal accounts) from 21 Yugan Khanty speakers, from which the locAg in active sentences can be found in 28 texts told (see Table 1). Besides the 19 excluded examples mentioned before, there are 54 other occurrences of the locAg in active sentences in the corpus – 3 with participles, 1 with a converb, 2 with infinitives and 48 with active finite verbs. Of these 48 examples with locAg in active sentences 7 are excluded due to being ambiguous for the following reasons: (a) the spoken and the standardized version of the example sentence are too different; (b) the end of the example sentence is not clear; (c) the example is a translation; (d) the storyteller speaks a Yugan Khanty variant with /t/ which allows the verbal ending to be in active and in passive as well, as the two endings coincide. That is why the research presented here is based on 41 examples from 11 speakers (see Table 1).

As can be seen in Table 1, around one-third of all texts contain a locAg in active sentences (28/86; 33%). Comparing the fourth and fifth columns shows that among the texts with the locAg in active sentences, it appears in two-thirds of them once (19/28; 68%). There are seven texts with two examples from four storytellers (7/28; 25%, from AIK, ENK, TMJ, VIU) and three texts with three examples from three partly different storytellers (3/28; 11%, from AIK, ANA, MDA).

Table 1: Summary of the texts in the corpus

	Storyteller	Texts told altogether	Texts containing locAg in active sentences	Number of examples with locAg active sentences
1	AIK[5]	11	5	8
2	AJK	1	0	0
3	AJM	4	0	0
4	ANA	1	1	3
5	ANB	1	1	1
6	AVJ	1	0	0
7	ENK	4	3	5
8	JFP	6	0	0
9	LDK	2	2	2
10	LNK	5	1	1
11	MDA	2	2	4
12	NEK	2	0	0
13	NPA	1	0	0
14	OAL	2	0	0
15	PDK	1	0	0
16	SPK	2	0	0
17	TAK	7	1	1
18	TMJ	11	4	5
19	TMK	5	1	1
20	VIU	16	7	10
21	VLK	1	0	0
	Sum	**86**	**28**	**41**

3. Assumed motivations for the use of the locative-marked agent in Yugan Khanty active sentences

In this pilot study, I will look into the question of the motivation of the locAg in active sentences in Yugan Khanty based on the corpus described above. The quantity of the data shows that this phenomenon in Yugan Khanty cannot be considered to be due to pure chance, mix-up, or speech disfluency, therefore the data will be analyzed from several sociolectal and grammatical aspects (**3.1–3.10.**). Not each point of the analysis will be elaborated in the same depth due to restrictions of the

[5] For the sake of anonymity, the storytellers are referred to with three-letter abbreviations.

data or the aim of this study. The data used in this paper is summed up in Table 12 in section **3.10.** Further summary of the findings is presented in section **5.**

3.1. Idiolectal motivation

The idiolectal motivation of the locAg in Yugan Khanty active sentences is suggested by the fact that one native speaker added or deleted the locAg 19 times while editing the manuscript of the storybook. But is it really an idiolectal phenomenon? Certainly not, as the native corrector is not the only person applying the locAg in active sentences. 52% of the storytellers (11/21) used the locAg in active sentences in the corpus. Out of the 11 speakers 6 used it more than once per text, and 5 speakers used it only once (see Table 1).

All eleven informants are native speakers of a local variant of Yugan Khanty, and were born in the vast and whole area of Yugan Khanty (for more details see **3.2.**). A majority of them are Khanty-Russian bilinguals (8/11; 73%), but there are three Khanty monolingual speakers (3/11; 27%). Most of them conduct a traditional way of life as fishers, hunters and gatherers in the forest along rivers (9/11; 82%), and few of them live in towns (2/11; 18%). Their age range is broad: the oldest speaker was born in 1932 and the youngest in 1969 (see Table 2). There are three male and eight female storytellers, so they are not uniform in gender, either. Some of them have no school education and are illiterate (2/11; 18%), some of them have few classes or are functional analphabets (2/11; 18%), some have up to eight classes (6/11; 55%) and one of them has even a university degree (1/11; 9%). Interestingly some storytellers were raised together (like ENK, LNK and partly TAK), some of them live together now (like ANA and MDA) and some have no direct, regular contact at all with the others (like ANB; see Table 2).

Despite the social-sociolectal differences between the speakers, all of them used the locAg in active sentences to some extent. A commonality of the 11 speakers explaining this phenomenon could not be found, so it is safe to say that the phenomenon under study is not motivated by idiolectal factors.

Table 2: Storytellers using locAg in active sentences

	Storyteller	Relatives	Born	Gender	Language
1	AIK	–	1962	female	monolingual
2	ANA	husband of MDA	1961	male	bilingual
3	ANB	–	1932	female	monolingual
4	ENK	sister of LNK cousin of TAK	1968	female	bilingual
5	LDK	son of TMJ	1969	male	bilingual
6	LNK	sister of ENK cousin of TAK	1966	female	bilingual
7	MDA	wife of ANA	1963	female	bilingual
8	TAK	cousin of ENK, LNK	1966	female	bilingual
9	TMJ	mother of LDK	1942	female	monolingual
10	TMK	–	1966	female	bilingual
11	VIU	–	1962	male	bilingual

3.2. Dialectal or subdialectal motivation

The question of the dialectal or subdialectal motivation of the locAg in active sentences arises, as this phenomenon is well attested in the Easternmost Khanty dialect Vakh-Vasyugan Khanty, which is in areal contact with Yugan Khanty. So, might it be an exclusive contact phenomenon only present in Yugan Khanty and non-existent in other Surgut Khanty subdialects?

To answer this question, we need to take a closer look at the geographical distribution of the Surgut Khanty subdialects. The focus of this paper is on the Yugan Khanty variant of Surgut Khanty which is the only southern subdialect of it, spoken along the rivers Bolshoy Yugan, Malyy Yugan, Salym and Demyanka in the 21st century, all situated geographically south of the river Ob. The other subdialects of Surgut Khanty, namely Pim, Tromagan and Agan Khanty, are spoken north of the Ob.

Going through the Surgut Khanty language materials available, it turns out that there are other examples for this phenomenon as well. There exist active sentences with a locAg in Tromagan Khanty in different collections of Márta Csepregi (1998, 2011; seven and three examples, respectively) or in my personal Tromagan Khanty archive (Schön 2016–2017; over fourteen examples). A local variant of Agan Khanty (Varyogan Khanty) contains at least two examples as well (Koshkareva 2006), and at a first glance, four examples could be found for Pim Khanty, too (Pesikova – Volkova 2013). Even if the numbers are relatively small, the existence of

the locAg in active sentences in all Surgut Khanty subdialects cannot be denied. It still has to be examined if it is a rare phenomenon or not, and its relation to the locAg in passive sentences is also worth looking at.

In addition to my Yugan Khanty corpus used in this paper, the locAg in Yugan Khanty active sentences can be found in other Yugan Khanty materials as well, like in Balalaeva et al. (2021). Their approximately 20 examples from five other speakers than mine will be analyzed at a later point in time, but this data is of importance as it shows that this phenomenon is not a new one as some texts were collected in 1994 from speakers born in 1920 or in 1936 (Balalaeva et al. 2021: 9–12, 136–138).

Furthermore, I want to emphasize that the locAg in Yugan Khanty active sentences is not a geographically restricted phenomenon as my storytellers come from the whole area and show mobility within it by comparing their place of birth and place of living (see Table 3).

Table 3: Mobility of storytellers using locAg

Storyteller	**Place of birth**	**Place of recording/living**
AIK	Demyanka[6]	Bolshoy Yugan (lower stream)
ANA	Bolshoy Yugan (lower stream)	Malyy Yugan (middle stream)
ANB	Malyy Yugan	Bolshoy Yugan (lower stream)
ENK	Salym	Malyy Yugan (upper stream)
LDK	Bolshoy Yugan	Bolshoy Yugan (upper stream)
LNK	Salym	Ob
MDA	Malyy Yugan (upper stream)	Malyy Yugan (middle stream)
TAK	Salym	Salym
TMJ	Bolshoy Yugan	Bolshoy Yugan (upper stream)
TMK	Malyy Yugan (upper stream)	Malyy Yugan (upper stream)
VIU	Bolshoy Yugan (middle stream)	Bolshoy Yugan (middle stream)

[6] For easier comparison of the local variants spoken, only the river names are given in Table 3 as places of birth and recording (identical to the place of living).

Even if it cannot be decided whether the origin of the locAg in active sentences goes back to contacts with Vasyugan Khanty (see also section **3.3.**), it is safe to say that it is attested in all subdialects of Surgut Khanty. Its presence might be observable to various extents, but it is not an exclusive Yugan Khanty phenomenon. Furthermore, Virtanen – Sosa (2018: 245) suggest that locAg in active sentences might appear in Surgut Khanty because of language contact with Tundra Nenets. To my knowledge, Tundra Nenets is in closer contact with Agan and Tromagan Khanty than with the other Surgut Khanty subdialects, but even this possibility would refute the restriction of the phenomenon to Yugan Khanty.

3.3. Influence from ritual language

The literature handling the locAg in Vakh-Vasyugan Khanty[7] clearly states that the locAg in active sentences often occurs in this dialect (Filchenko 2006: 57, Csepregi 2020: 45), even in mythological or ritual songs (Csepregi 2020). It could be assumed that this influenced its appearance in Yugan Khanty. There exists one study about the language of Surgut Khanty languelteps: the work of Márta Csepregi with the material of K. F. Karjalainen, where only one example of the locAg in active sentences was found (Csepregi 2020: 45–46).

Even if the study presented in this paper deals with narratives (tales, legends and personal accounts) that have a different, more open structure than ritual songs and are more similar to spoken language, let's follow the supposition that if the storytellers have a deep understanding of the language of mythological songs, they could use it in narratives, too.

It is without a doubt that all eleven informants have an elaborate knowledge of Yugan Khanty folklore, inherited from their ancestors, relatives or learnt from their environment. This knowledge is transmitted orally. In addition to this and the language skills, the speakers have to have a good memory as well as the ability of storytelling, singing too. Traditionally, only men are allowed to sing most of the mythological songs, but women are allowed to listen to most of them (personal account from Stephan Dudeck 2009). So, a differentiation between the active and passive knowledge of the language of ritual songs has to be made (see Table 4).

[7] The same as Eastern Khanty in Filchenko (2006, 2011).

Table 4: Knowledge of ritual songs of the storytellers[8]

Story-teller	Gender	Could sing Surgut Khanty languelteps	Has heard Surgut Khanty languelteps	Has heard Vakh-Vasyugan Khanty languelteps
AIK	female	no	no	no
ANA	male	yes	yes	perhaps
ANB	female	no	yes	no
ENK	female	no	no	no
LDK	male	yes	yes	perhaps
LNK	female	yes	yes	no
MDA	female	no	yes	perhaps
TAK	female	no	no	no
TMJ	female	yes	yes	perhaps
TMK	female	yes	yes	no
VIU	male	yes	yes	perhaps

The fact that most of my storytellers are women (8/11; 73%) does not exclude the possibility of this influence. It is rather the heterogeneity present among them here as well: around half of them could sing ritual songs (6/11; 55%), some have heard them regularly (8/11; 73%) but others are presumed not to (3/11; 27%; see Table 4). Furthermore, the fact that, till today, the occurrence of locAg in active sentences in Surgut Khanty ritual songs is not really attested, speaks also against its source being those songs. In my personal experience, none of my informants could have heard Vakh-Vasyugan Khanty ritual songs regularly. Some of them clearly stated that when hearing Vasyugan Khanty people speaking Vasyugan Khanty, they did not understand them and switched to Russian with them (personal account from beside the Bolshoy Yugan river, 2015).

3.4. Dependence on verbal semantics

The question if the appearance of the locAg in Yugan Khanty active sentences depends on verbal semantics is based on typological features and the fact that in some languages only certain verbs can cooccur with a marked subject or agent. It is known about Icelandic that marked sub-

[8] The information in Table 4 is based on my personal experience and observations while working with the storytellers, as well as on my deduction from their curricula.

jects or "non-nominative subjects" only occur with verbs of restricted or special semantics, and that is known to be the case in German, too (Sigurðsson 2007).

The verbal semantics of the 41 Yugan Khanty examples with a locAg in active sentences in the corpus shows a high degree of variation. They contain 30 different verbs with 25 different meanings, as shown in Table 5.

Table 5: Verbal semantics of the active verbs with locAg in Yugan Khanty active sentences

Gloss	Verb	Occurrences	Gloss	Verb	Occurrences
'do'	*βɛr-*[9]	2	'string'	*kɛrit-*	1
'give'	*mə-*	3	'engrave'	*tʃomәłtə-*	2
'send'	*kiːt-*	2	'have'	*taj-*	1
'bring'	*tuː-* *ɐːł-*	1 1	'cheat'	*łɐːjɣәł-*	1
'put'	*pan-*	1	'see'	*βuːjitɛɣrәɣtə-*	1
'come'	*joβәt-*	3	'think up'	*nomłәmtə-*	1
'throw'	*jɛβәt-*	2	'allow'	*ɛsәł-*	1
'raise'	*noq kiːłtə-*	1	'summon'	*βɒːɣ-*	1
'close (eye)'	*ɯːłə qonʲ-*	3	'invite'	*βɒːɣił-*	1
'poke out'	*jɐːɣli-* *pɯːɣәmtə-*	1 1	'say'	*βɛr-* *kәłijłə-*	2 1
'stretch out (leg)'	*nʉrtәmtə-* *iːłə nʉrt-*	1 1	'turn to smb and say'	*jiɣ-*	1
'break'	*iːłə* *moːrimtə-*	1	'ask'	*pɯːripə-* *pɯːrił-*	1 1
'chop into pieces'	*iːłə kөsə-*	1			

A commonality or an anchoring point of the verbs is difficult to find, which could suggest productivity. Most of them are perfective verbs or used in a perfective way, but not all ('do', 'come'). There are verbs of diction ('say', 'ask'), verbs of motion ('come'), and body part manipulation ('close (eye)', 'stretch out (leg)'). Most of the verbs are active, their locAgs can be more or less in control but there is only one verb with an experiencer-type subject (4) and even an intransitive one occurs in this construction (7c), (8), (9c).

[9] The verbs in Table 5 are given with verbal stems.

(4)	***tʲuː***	***qur***	***mɐːn***	*pə*	***βuːjitɛɣrəɣt-em***
	that	**treeless_marsh**	**1SG.LOC**	EMPH	**see.PST-SG<1SG**

'I have seen that treeless marsh.' (Kayukova – Schön 2017)

In example (4)[10] the storyteller herself gives a comment about a place in a legend to underline its existence at the end of the story (sentence 164/170). Interruptions and questions are characteristic of this narrative, as the storyteller frequently turns to her listeners with questions and comments, but this is the only occurrence when AIK refers to herself. The verb itself seems to be her own creation as it contains several derivational suffixes – its stem is *βuː-* (DEWOS 1551).

Having seen the diverse nature and broad semantics of the verbs appearing in active voice with a locAg, it can be said that Yugan Khanty does not have a restricted set of verbs occurring with a marked subject, namely the locAg, in active sentences, so one similarity with the "non-nominative subjects" can be excluded.

3.5. Dependence on transitivity or conjugation

Surgut Khanty, including Yugan Khanty, has intransitive and transitive verbs, as well as a subjective and an objective conjugation, where the verb shows agreement either with the (grammatical) subject or with the (grammatical) subject and the direct object (DO). The dependence on transitivity of the locAg in Khanty active sentences is relevant because opinions differ as to whether the construction occurs exclusively with transitive verbs (Csepregi 2021) or with both transitive and intransitive verbs (Sosa 2017: iv). Concerning the issue of conjugation, it is important to note that in Khanty if the DO is topical then objective conjugation appears, and subjective conjugation shows if it is not (Nikolaeva 1999: 365). As the locAg is often said to appear in the case of two concurring topics (Sosa 2017: 191–207), the choice of conjugation could play a role in its occurrence as well.

In my corpus, the locAg in active sentences can appear both with transitive and intransitive verbs in subjective or objective conjugation (see Table 6).

[10] Further details concerning this example are provided in section **3.7.**

Table 6: Correlation between transitivity and conjugation of the verbs examined

	Subjective conjugation	**Objective conjugation**
Transitive verb	10 examples	28 examples
Intransitive verb	3 examples	–

As shown in Table 6, transitive verbs in objective conjugation (3), (4) are more commonly used in the corpus (28/41; 68%), but intriguingly, about one-third of the examples (13/41; 32%) contains verbs in subjective conjugation.

The examples with transitive verbs in subjective conjugation (10/41; 25%) show large variation as only one verb ('do') occurs twice, so the ten examples are formed with nine different verbs and semantics. They are heterogeneous considering the DO, too: five examples contain no DO at all (5), in three examples the lexically overt DO is expressed with a noun (15), (22g), and in two with a personal pronoun in the accusative (6).

(5) ***tɯβnə*** ***jiɣ-it*** *oːs* *mɐː*
3SG.LOC **turn_to_smb_and_say-PRS.3SG** and 1SG
nɯŋɐt *mətt* *qotnə* *ojɐɣtə-t-ɛm* *məttə*
2SG.ACC (s)he_says how find-PRS-SG<1SG (s)he_says
nɯŋ *liːk-ən* *qotnə* *ojɐɣtə-t-ɛm*
2SG path-SG<2SG how find-PRS-SG<1SG

'He turns around and says: "And how will I – he says – find you – he says – how will I find your path?"'

(Schön 2016a)[11]

(6) *jɐː* *oːs* ***teːmi*** ***iːmi-liŋki-nə*** ***mɐːnt***
well again **this_here** **wife-DIM.PEJ-LOC** **1SG.ACC**
kiːt *ənət* *t̮orɐs* *qoː* *kiːm-ɐ* *βɛr-tɐː*
send.PST.3SG big merchant man possibility-LAT do-INF

'Well, this poor wife [of mine] has sent me again, so that you make us rich merchants.' (Schön 2019a)

[11] I recorded this plot from the same storyteller (TMK) twice with three days' difference (24.6.2012 and 27.06.2012). In the first version of the tale, there is no locAg in the same sentence (see sentence number 38 in Schön 2017a), it appears only in the second version as in (5).

Even though intransitive verbs in subjective conjugation are rare (3/41; 7%), they are the most interesting. In all three examples (7–9), the same verb ('come') is used, but in three different contexts.

In (7a–c) we see the first three sentences of a tale. The locAg appears in the third sentence as a threat (7c), after the frame of the tale is set: a woman is awake working nights, when everybody else in the room is already asleep, even if this is forbidden (7a–b). This introduction of the plot is common knowledge in the Khanty worldview. The presence of the locAg at the beginning highlights its saliency and underlines the teaching function of tales. I have elicited this plot two more times from two different storytellers (TMK, SPK – the first of whom used the locAg in other narratives, the second never), and in contrast, in both their versions the threat itself is missing completely (see Schön 2016b; Kayukova – Schön 2016a).

(7a)	*əj*	*βəsyə*	*marə*	*ɒ:məs*		*ij*	*ne:*
	once	well	time	sit.PST.3SG		one	woman

'Once a woman sat [awake] for a long time.'

(7b)	*qoβ*	*mar*	*əj*	*ɒ:ł*	*βiłɛ*	*əntə*	*ɒ:məs-t-i*
	long	time	one	place	in_fact	NEG	sit-PRS-PASS.3SG

'One doesn't sit [awake] for a long time.'

(7c)	***jipəɣ-nə***	***joɣot-ł***
	fright-LOC	**come-PRS.3SG**

'[Or else] the monster is coming.' (Kayukova – Schön 2016b)

In example (8), the locAg in an active sentence occurs in the middle of the text. It emphasizes the fact of what happens if someone breaks the (language) taboo rules. In the plot, it is forbidden to say the name of the sacred river (*mɛŋkət jɐ:ɣ* 'Menks River'), otherwise something bad happens – a Menk, a monster comes. The protagonist breaking this taboo is a stranger, a relative-in-law, namely the Son-in-law from the Ob. He is told about this language taboo in the plot by his wife's younger brother, so we can exclude the idea of common knowledge (present in (7a–c)). In the second version of this tale, which I collected from another storyteller (VIU), the situation is the same, but the concrete consequence of the taboo breaking is missing as well as the locAg in the similar constructions (Schön 2017b).

(8) *mɛŋk-ət* *jɐːɣə* *jɐːstə-t-i* *quːntə* *pɐːn*
Menk-PL river say-PRS-PASS.3SG if then

mɛŋk-nə ***joɣot-ł***
Menk-LOC **come-PRS.3SG**

‘If one says Menks River, then the Menk comes.’
(Kayukova – Schön 2017)

Even if examples (7a–c) and (8) are from the same storyteller (AIK) and the locAg in both examples could be interpreted similarly because of breaking a behavioral (7c) or language taboo (8) rule, in example (9) we find an entirely different situation in a third plot from a personal account and a second storyteller (TMJ).

(9a) *ӘЙ* *ӆат-нə* *пətəӽӆ-əм* *пырнə*
one time-LOC darken-PTCP.PST after

мəта-нə *өвпи* *тьи* *пунч-и*
something-LOC door so open.PST-PASS.3SG

‘Once, after it got dark, something opened the door.’

(9b) *Атьи-ӆ-а* *пəтьа* *вӓр*
father-SG<3SG-LAT hello say.PST.3SG

‘It said hello to the father.’

(9c) *Аӈкэ-м* *ньӑвмиӆ-əӆ* *Та*
mother-SG<1SG begin_to_speak-PRS.3SG look

тыӽəӆ *та* ***Ньəмньан*** ***ики-нə*** ***йӧвəт***
(s)he_says look **Demyanka** **man-LOC** **come.PST.3SG**

‘My mother begins to say: “Look!” – she says – “Look! Uncle-Demyanka came!”’ (Schön 2015b)

The referent of the locAg in the active sentence (9c) is already introduced with a pronoun (*мəтанə* ‘by something’) in the first sentence (9a) as the locAg in a passive sentence. It appears in the second sentence too as an unmarked 3SG verbal ending (9b) and once more with a proper noun (*Ньəмньан икинə* ‘(by) Uncle-Demyanka’) in the direct speech of another protagonist (*Аӈкэм* ‘my mother’; 9c). As the person of the locAg in the active sentence is only identified in its third occurrence, the context is highly emphatic, one could even think about the locAg in the active sentence (9c) as the expression of mirativity.

Altogether my data shows that the locAg in active sentences can occur both with transitive (38/41; 93%) and intransitive verbs (3/41; 7%), so the transitivity of the verb is not essential in the appearance of the locAg in active sentences. Concerning conjugation, one third of the examples show only subject agreement (subjective conjugation; 13/41; 32%) and two thirds of them show subject-object agreement (objective conjugation; 28/41; 68%), so conjugation is not an exclusive factor for the use of locAg in active sentences, either. Additionally, the data attests that there are not always two concurring topics from which one (the subject/agent) has to be marked with the locative case to be distinguished from the other (the DO) as there is no accusative case for nouns in Khanty. The question of topicality definitely requires further research.

3.6. Correlation with word order

The neutral word order in Surgut Khanty is SOV (Honti 1984: 88), where S is dominantly a grammatical subject (1). This is also the case in the majority of the sentences cited here, that is, in active sentences with a locAg (3–9). I restrict the object (O) to a direct object (DO) in this paper, and mean by V a finite verb in active conjugation. A change of this neutral SOV word order can be caused by Russian contact influence (Gugán – Sipos 2017: 77) or explained by the subject as a known referent (Asztalos – Gugán – Mus 2017: 30–31). The phenomenon of afterthought can also be an explanation for a non-neutral word order. But can word order or even word order change be correlated to the locAg in Yugan Khanty active sentences?

My data indicates that S(O)V word order in sentences with locAg and an active verb is the most common (33/41; 80%), but there are examples for OSV (2/41; 5%) and for (O)VS (6/41; 15%) as well (see Table 7).

Table 7: Word order and conjugation in Yugan Khanty examples with locAg and active sentences

	Subjective conjugation	**Objective conjugation**
SOV	5 examples	13 examples
OSV	–	2 examples
OVS	–	5 examples
SV	8 examples	7 examples
VS	–	1 example

Some word orders were already shown, like SOV and subjective conjugation (6), SV and subjective conjugation (5), (7c), (8), (9c) or objective

conjugation (3), or OSV and objective conjugation (4). In example (10) a neutral SOV word order is given, with a personal pronoun as locAg (S; *нӱԓнə* ‘(by) you’), a noun phrase as DO (*əй нӗв* ‘one branch’) and a verb with preverb and objective conjugation (*илԓə кӧсэ* ‘you chopped it into pieces’).

(10) *Панə* ***нӱԓнə*** ***əй*** ***нӗв*** ***илԓə***
and **2SG.LOC** **one** **branch** **PFV**
кӧс-э *мөримт-э*
chop_into_pieces.PST-SG<2SG break.PST-SG<2SG

‘And you chopped one branch into pieces, you broke it.’
(Schön 2023)

There is only one example for VS (11), where the locAg expressed with a personal pronoun (*ɬuβonə* ‘he/by him’) follows the verb (*puːrittətəɣ* ‘he asks her’). The reason for the position of the subject is supposed to be that the storyteller wanted to clarify that it is the man who is asking the female bear why she is still sitting with him, yet it does not bear the hallmarks of afterthoughts. Even if, grammatically, the 3SG.LOC could refer to the man as well as to the female bear, it is clear that it refers to the man, as the bear cannot talk in this plot, so she cannot be the subject of this sentence. The storyteller (MDA) is bilingual, so Russian influence cannot be excluded.

(11) *tʲuː* *ɒːməs-t-in-nə* *pɐːn* *əj* *tɐːt-nə*
so sit-PTCP.PRS-3DU-LOC then one time-LOC
tʲɛl ***puːrit-t-ətəɣ*** ***ɬuβonə*** *toɣən* *βəs* *ij* *riːs*
always **ask-PRS-SG<3SG** **3SG.LOC** so or one time

‘While they sit like that, then one day, he always asks her, many times, one time.’ (Schön 2017c)

Even if the word order OVS is not typical for Yugan Khanty, the five examples (5/41; 12%) from five different storytellers provide three different explanations. In one case (a) it appears in a ditransitive construction (12) clarifying the agent of the sentence pragmatically, because there is another “old man” in the previous sentence and (b) in two cases to emphasize even more who the agent of the sentence was (13), where in one case even a strong swear word (*tʲiː jimpəʃ otnə* ‘(by) this poor little soul’; VIU) is used. In two cases (c) it occurs as an afterthought (14).

Interestingly, three of the informants using the word order OVS with locAg in active sentences are monolingual (ANB (14), TMJ (12), AIK), and only two are bilingual (ENK (13), VIU). The two cases of afterthought are from the monolingual speakers, and the assumed emphasis as postverbal focus from bilingual speakers. This can be identified as a language contact phenomenon.

(12) *Панə* ***тьу*** ***йасӈ-ət*** *ӌӗмл-əӈ* *йуӽ-а*
and **that** **told-PL** notch-PROPR stick-LAT

ӌӗмəлт-əлəл ***тьу*** ***ики-нə***
engrave.PST-PL<3SG **that** **old_man-LOC**

'And that old man engraved all things told on the notched stick.'
(Schön 2015a)

(13) ***Ма*** ***пиӈ-əм*** ***илə*** ***мəримт-э*** ***нӱӈнə***
1SG **thumb-SG<1SG** **PFV** **break.PST-SG<2SG** **2SG.LOC**

'You broke up my thumb. / It was you who broke up my thumb.'
(Schön 2023)

(14) *tʲuː* *ɬɯβ* *tʲuː* *mɯŋətʲɣiɬtə-t-ɐɬ-ən* *pɐːnə*
so 3SG so go_and_loom-PTCP.PRS-3SG-LOC then

kur-əɬ *toɣə* ***nurtəmt-ətəɣ*** *tʲuː* *…*
leg-SG<3SG there **stretch_out.PST-SG<3SG** that **...**

tʲuː ***toɣ*** ***alint-əm*** ***iːki-nə***
that **there** **lie_down-PTCP.PST** **man-LOC**

'So, while she left and loomed, he stretched out his leg there, that... that man, who had lain down there.'
(Kayukova – Schön 2020a)

Example (14) is the turning point of the plot: the female ghost, looming around, will trip and fall because of the legs stretched out across her path by the mortal man. Listening to the recording,[12] it is clearly audible that the monolingual, elderly speaker ANB adds the lexically overt locAg as S to the sentence with a longer pause, so this additional explanation could be interpreted even as a separate sentence. The motivation for adding an afterthought is the fact that in (14) there are two competing refer-

[12] Available at https://oudb.gwi.uni-muenchen.de/?abfrage=view_audio_files&id_text=1658

ents marked both with 3SG (1. she, the female ghost marked with the 3SG personal pronoun *łuβ* and the 3SG ending on the participle and 2. he, the male human protagonist marked with the 3SG verbal ending) and there is a topic switch between them, so the second referent should normally be marked lexically.

Asztalos et al. state that in old Surgut Khanty texts (1901, Paasonen's collection) only those subjects occurred postverbally whose referents were introduced earlier (Asztalos et al. 2017: 33, 44; Gugán – Sipos 2017: 90), but in contemporary Surgut Khanty (recordings from 1994, 2000, 2004) also newly introduced referents may occur in a VS position (Asztalos et al. 2017: 34, 46). My data (2010–2015) show similarities with the old Surgut Khanty data. In all six examples for (O)VS, the locAg in verb-final position expresses old information, as all referents/locAgs were already introduced somehow in the texts before, which can be tales (11), (13), legends (14), or personal accounts (12). This is perhaps due to the fact that the "old data" from 1901 consists of tales and legends, i.e., folklore material, as does most of my corpus, and the contemporary Surgut Khanty texts represent personal accounts.

After taking a closer look at the examples with VS (11) and OVS (12), (13), (14) word orders, it has to be recalled that the majority of my data (33/41; 80%) contains the neutral SOV word order (10). So even if some word orders can be explained by Russian influence or afterthought, it is clear that word order or word order changes are not the motivating factors behind the use of the locAg in Yugan Khanty active sentences.

3.7. The internal structure of the subject in Surgut Khanty as a locative-marked agent in active sentences

All Khanty dialects are dominantly nominative languages, so in a standard sentence the subject is in the nominative case (if lexically present) and always shows agreement with the verb. A divergence in the form of the lexically overt subject can be observed: it can be a common noun (with or without modifiers), a proper noun, an adjective, a numeral, or different pronouns. Khanty being a pro-drop language, the subject (and the direct object) can exclusively be grammatically overt. Compared to this, if the subject is expressed with the locAg in an active sentence, the locAg is always lexically overt.

Sosa states that the locAg in Khanty active sentences is generally expressed by personal pronouns (Sosa 2017: 185, 191). Filchenko finds that it can be "expressed by a free pronominal NP or a full NP" (Fil-

chenko 2006: 61), but in 71% of his data the locAg in an active sentence is a personal pronoun (Filchenko 2006: 61). But then, does the locAg in active sentences have to be mostly expressed by "inherently definite expressions (such as proper nouns, personal and demonstrative pronouns)" (Klumpp – Skribnik 2022: 1017) in order to be accessible or identifiable?

At first glance, one might assume that the form of the lexically overt locAg in active sentences in my corpus (see Table 8) is diverse. Taking a closer look, however, it becomes clear that in 76% (31/41) of the data the locAg is expressed by a common noun (with or without modifiers) or a proper noun, and in 24% (10/41) by a pronoun, dominantly a personal pronoun in the locative case (9/10; 90%).

Table 8: Expressions of the locAg in Yugan Khanty active sentences of the corpus[13]

	Occurrence	**Examples cited**
ppron	9 examples (1SG.LOC 2, 2SG.LOC 3, 3SG.LOC 4)	(3), (4), (5), (10), (11) (13), (15b)
nprop	2 examples	(9c), (20)
dem+subs[14]**-POSS**	2 examples (SG<1SG, SG<3SG)	–
dem+subs	16 examples	(6), (12), (18), (19), (22g)
dem+adj+subs	1 example	–
dem+adv+ptcp+subs	1 example	(14)
subs-POSS	1 example (SG<1SG)	(17)
subs	8 examples	(7c), (8), (21f)
interrog	1 example	–

Unfortunately, not all personal pronouns constitute for sure the locAg of an active verb. There exists a homonymy between the 1SG locative personal pronoun (*mɐ:nə* '(by) me') and the short form of the 1SG emphatic personal pronoun (*mɐ:nə* 'myself'). The 1SG emphatic personal pronoun often has the form *mɐ: mɐ:nə* 'I myself', but the first part (*mɐ:* 'I') can be

[13] The use of POS-tags in the tables follows the tagging system of the OUDB corpus.

[14] Including compounds like *torəm qɒ:n* 'tsar'.

omitted, so it has only the form *mɐ:nə* ‘myself’, identical to *mɐ:nə* ‘I, (by) me, 1SG.LOC’ (see Table 9 for the whole paradigms).

Table 9: Paradigms of the locative and emphatic personal pronouns

	Locative personal pronoun	**Emphatic personal pronoun**
1SG	*mɐ:nə*	*(mɐ:) mɐ:nə*
2SG	*nʉŋnə*	*(nʉŋ) nʉŋə*
3SG	*łʉβnə*	*(łʉβ) łʉβə*
1DU	*mi:nɐtemnə*	*mi:nmin*
2DU	*ni:nɐtinnə*	*?*
3DU	*łi:nɐtinnə*	*łi:n łi:nə*
1PL	*məŋnə*	*məŋ məŋə*
2PL	*nəŋɐtinnə*	*?*
3PL	*łəɣnə*	*łəɣ łəɣə*
	(Csepregi 1998a: 24–25)	(Yugan Khanty corpus)

Hence, all examples containing *mɐ:nə* ‘1SG.LOC’, ‘by me’, ‘I’ or ‘(I) myself’ are at least ambiguous.[15]

Apparently, some data presented in the literature are also problematic for the very same reason. Certain sentences published in Sosa (2017) or Sosa – Virtanen (2018) are at least ambiguous or even misinterpreted owing to the homonymy of the 1SG.LOC personal pronoun and the emphatic personal pronoun *mɐ:nə* ‘I myself’ in 1SG,[16] which Sosa did not take into account.

In example (6.74 a) *manə nürəɣətəm tom jäŋk päləka* ‘I ran to the other side of the small lake.’ (Sosa 2017: 184, citing Csepregi 1998a: 60), the presumed locAg *manə/mɐ:nə* was corrected during the editing of Csepregi’s collection (1998a) in the OUDB project in 2011 by Csepregi herself into *mɐ:nɐm* ‘myself’ (Csepregi 1998b), which is a second form of the 1SG

[15] There are two such examples in my corpus: (4) and (15b).

[16] Taking a closer look at the eight examples cited by Sosa with a personal pronoun as locAg in active sentences, it can be seen that five examples are with a form of *łʉβnə* ‘3SG.LOC’, ‘he/she, (by) him/her’ and three with a realization of *mɐ:nə* ‘1SG.LOC’, ‘(by) me’ (Sosa 2017: 184, 189–191), which are at least ambiguous, because of the homonymy with the emphatic personal pronoun *mɐ:nə* ‘I myself’ in 1SG. In the data from my corpus used in this paper, there are altogether 9 examples (out of 41) for the locAg with active sentences expressed by a personal pronoun, namely twice with *mɐ:nə* ‘I, (by) me, 1SG.LOC’ (4), (15b), three times with *nʉŋnə* ‘(by) you, 2SG.LOC’ (3), (10), (13) and four times with *łʉβnə* ‘he/she, (by) him/her, 3SG.LOC’ (5), (11).

emphatic personal pronoun. Listening to the sound recording proves that the speaker pronounces *mɐːnɐm* 'myself' instead of *mɐːnə* '1SG.LOC' or 'myself' (*pɐːn mɐːnɐm nʉrəɣtəm tom jɛŋk pɛləkɐ* 00:37–00:40[17] Csepregi 1998b). This could help with the choice in the case of the homonymy of *mɐːnə*, or even mean that this sentence does not contain the locAg anymore (as in the second edition of the tale, Csepregi 1998b).

In another example used by Sosa, namely (6.74 b) *manə waləɣ qŏwit utnam quɣləm* (2017: 184, citing Csepregi 1998a: 62), the correct interpretation of the verb itself and the whole sentence is doubtful, in addition to the ambiguity of *manə/mɐːnə* '1SG.LOC' or 'myself'. Sosa and I earlier[18] had identified the verbal stem *quɣl-/quːɣł-* as a past stem of *qoːβəł-* 'run' with paradigmatic vowel change, but consulting the Hungarian translation of Csepregi (1998a: 63, 1998c) and listening to the sound recording (00:55–01:01)[19] raised serious doubt. One alternative interpretation of this example could be that the informant says *quːŋt-* 'climb' (DEWOS 528, SVH 145), which would explain the Hungarian translation of Csepregi ('én meg a rénhajtó rúd segítségével másztam ki' [and I climbed out with the help of the driving-rod] Csepregi 1998c; English translation by Katalin Gugán). This presumption is supported, as on the sound recording instead of the *-l-/-ł-* in the verbal stem a *-t-* is clearly audible. There is a second alternative to interpret example (6.74 b): the verb *quɣl-/quːɣł-* could be a different verb with the meaning 'zur Wasserfläche emporsteigen, auftauchen', 'zur Oberfläche (des Wassers, des Kessels) emporsteigen, auftauchen' from the DEWOS (454 and 455, respectively), but unfortunately, there are no Surgut Khanty verbal forms attested, just from other Khanty dialects. Anyhow, the interpretation of this example (6.74 b) given by Sosa ('I ran to the shore after the bar of the reindeer sled.' 2017: 184) cannot be valid, as the plot goes in the sentences before like this: the storyteller traveling on a reindeer sledge in winter time falls into the water after the ice crushed beneath him. He cannot reach the shore and is only able to climb out of the water with the help of (in the text the word 'along' is used) the reindeer driving-rod. Interestingly, in the sentences coming before and after this example used

[17] Audio recording available at https://oudb.gwi.uni-muenchen.de/?abfrage=view_audio_files&id_text=728

[18] In Csepregi (1998c) the glosses were made by me.

[19] Audio recording available at https://oudb.gwi.uni-muenchen.de/?abfrage=view_audio_files&id_text=730

by Sosa two verbs in passive appear and the corresponding locAgs are lexically overt (*mɐːn* 'by me', *mɐːnə* 'by me' Csepregi 1998c), so even if I agreed with Sosa on the verb and interpretation of the example, the possibility of priming cannot be excluded either. Sosa has taken a look at the textual surroundings of the example, but has come to different conclusions (2017: 189).[20]

Another indicator to determine the meaning of *mɐːnə* can be the conjugation of the corresponding active verb – if the objective conjugation appears, then *mɐːnə* is more likely a locAg (4?).[21] A second hint can be the presence of particles, like the emphatic particle (*pi*/*pə*) in example (15b?).

(15a)	*mɐː*	*βitɛ*	*əjjɐːk-kən*	*taj-t-əm*	*nɛritnə*
	1SG	in_fact	parents-DU	have-PRS-1SG	[unknown]

'In fact, I have parents.'

(15b)	***mɐːn***[22]	*pi*	***jaq***	***tɐːɣ***	***taj-t-əm***
	1SG.LOC//1SG.EMPH	EMPH	**home**	**place**	**have-PRS-1SG**

'I do have a homeland.' // 'I have a homeland myself too.'

(Schön 2017d)

Although in example (15b?) we have an emphatic particle (*pi* 'EMPH') and a verb in subjective conjugation (*tajtəm* 'I have'), it is hard to decide if *mɐːnə* has to be interpreted as locAg or as an emphatic pronoun. In my opinion it can be either.

[20] Unfortunately, these two examples found by Sosa (6.74 a and 6.74 b; 2017: 184) are the only ones in her dissertation that support the existence of intransitive verbs in Surgut Khanty active sentences with locAg. So, her statement of "the locative subject of an intransitive verb is always primarily topical" (2017: 189) is at least problematic as well, but my research proves the existence of locAg in Surgut Khanty with active and intransitive verbs with more convincing examples, as shown in (7c), (8), (9c).

[21] Because of the homonymy between the locAg expressed by a 1SG personal pronoun and a 1SG emphatic personal pronoun, the examples (4) and (15b) from now on are marked with a question mark (4?), (15b?) and remain part of the 41 examples used in this paper as such.

[22] The reduced vowel at the end of *mɐːnə* can be dropped, regardless of its meaning.

In the oldest Yugan Khanty text collection from Heikki Paasonen only one presumed example of the locAg in an active sentence can be found, with *mɐːnə* (16).[23]

(16) *iːmi* *jɐːstə-ł* ***mɐːn*** *pɐ* *łiː-ttɐːə* ***jəɣɣ-əm***
woman say-PRS.3SG **1SG.LOC** too eat-INF **become.PST-1SG**
'The woman says: "I've gotten hungry, too."'
(Paasonen – Vértes 2001)

It has to be pointed out that in my corpus the locAg in Khanty active sentences can be expressed by a personal pronoun, a proper noun, as well as by a common noun (with or without a modifier), all in the locative case. Interestingly, the noun phrases in my data show a high proportion of the combination of demonstratives and nouns in the locative (20/41; 49%) that are present twice as often as pronouns in the locative (10/41; 24%). Non-modified nouns in the locative alone appear almost to the same extent as pronouns do (8/41; 20%), and there is only one locAg expressed with a possessive-marked noun in the locative (1/41; 3%). Fortunately, only two examples with the 1SG.LOC personal pronoun are ambiguous in my corpus because of the homonymy with the 1SG.EMPH. Most of the personal pronouns as locAgs are in other persons (2SG.LOC 3 examples, 3SG.LOC 4 examples). Altogether, my data shows that the locAg mostly has to be accessible or identifiable already by its grammatical form (9+2+20+1+1/41; 80%), but a non-modified noun in the locative can appear as locAg as well (8/41; 20%).

3.8. The internal structure of the direct object in active sentences with a locative-marked agent

The direct object (DO) in Khanty is often unmarked, as there is no accusative case marker for nouns in Khanty; only personal pronouns have an accusative form.

Klumpp and Skribnik state that the locAg in East Khanty active sentences "is used if both, the subject and the object, are more or less equally topical, but the patient is superior in animacy (often this is a personal pronoun)" (Klumpp – Skribnik 2022: 1024). Contrary to this, Sosa finds

[23] Paasonen collected his Yugan Khanty material in 1901 and it was first published in 2001 by Edith Vértes. Here the example is cited in another version from the project OUDB, referred to as Paasonen – Vértes (2001).

that if the verb is transitive, then the "predominate referential form of an object with a locative subject is affixal: three [examples] are affixal (object conjugation), three are LEX+V (object conjugation plus an overt lexical utterance) and one is a lexical utterance (with the subject conjugation)" (Sosa 2017: 191). Filchenko finds in his data that the object, if it is present in a sentence, is an identifiable NP (Filchenko 2006: 61). The question arises whether a DO is obligatorily present in every construction with locAg in active sentences, and whether the locative therefore appears because of the topicality of the S and the DO at the same time. Or is the DO not necessary with the locAg in active sentences? And finally, if a DO is present, which internal structure does it have to take in order to be identifiable?

Before I present the examples with grammatically or lexically overt DO in my corpus, I want to analyze the examples without any DO, because in one-fifth of the examples of this study there is no DO at all (8/41; 20%). In all these 8 cases the verb is in subjective conjugation (see Table 10). We have already seen three examples of no DO with the intransitive verb 'come' (7c), (8), (9c). In example (5) we saw a transitive verb in subjective conjugation (*jiɣ-* 'turn to somebody and say') and no DO, but here the locAg construction is followed by direct speech as typical for *verba dicendi* or *verba intelligendi*. There is a third type of examples for locAg in active sentences without any DO as demonstrated in example (17). Therefore, it is quite clear that the presence of two topics or a DO cannot be the trigger for locAgs in active sentences in every case.

(17) *əj* *tɐːt-nə* *noməqsəɣə-t* *oːs*
one time-LOC think-PRS.3SG lo_and_behold
iːmɛ-m-nə *otəŋ* *ottə* *tʲeːm* *qpːt*
aunty-SG<1SG-LOC whether er this house
potʃ-əŋ *pɛləkˑnɐm* *mɯβɐt* *ənt* ***ɛsət-t***
back-PROPR side-APP why NEG **allow-PRS.3SG**

'One day he thinks: lo and behold! my aunty, er, why doesn't she allow me to go to the back side of this house?'

(Schön 2019b)

The data of my corpus indicates that the DO in the active sentences with locAg can be expressed lexically, grammatically, as well as lexically and grammatically at the same time (see Table 10). If the DO is only lexically

overt, it stands with a verb in subjective conjugation (5/41; 12%; (6)), if it is only grammatically overt, it stands with objective conjugation (8/41; 20%; (3), (11)). Objective conjugation appears if the DO is topical. If the verb shows only subject agreement, then the DO is supposed to be focal.

Table 10: Correlation of conjugation and the direct object in the Yugan Khanty active sentences with locative-marked agent[24]

	Subjective conjugation	**Objective conjugation**
Lexically overt direct object	5 examples	–
Lexically + grammatically overt direct object	–	21 examples
Grammatically overt direct object	–	8 examples
No direct object	8 examples	–

The DO is grammatically overt if the verbal ending shows object agreement. One-fifth of my data contains only a grammatically overt object (8/41; 20%) as shown in examples (3), (11).

In terms of their internal structure, the lexically and the lexically and grammatically overt DOs show large variation at first sight (see Table 11). Taking a closer look, it turns out that almost in all examples containing a DO, this is expressed with some kind of a noun (23/26; 88%; (4?), (10), (12–14), (15b?)) and only in a few cases with pronouns (3/26; 12%; (6)). This is contraindicative to the DO (patient) being often expressed by personal pronouns and so being superior in animacy to the locAg (*pace* Klumpp – Skribnik 2022: 1024).

[24] One example contains two DO objects (20), which is counted twice in that case, that is why here it seems if the corpus contained 42 examples, instead of 41.

Table 11: Expressions of lexically or lexically and grammatically overt DO in Yugan Khanty active sentences with locAg

	Lexically overt direct object	**Lexically + grammatically overt direct object**
ppron	2 examples (1SG.ACC, 1PL.ACC)	–
dem-POSS	–	1 example (PL<3SG)
subs	2 examples	2 examples
subs-POSS	1 example (SG<3SG)	8 examples (PL<1SG; 4 DU<3SG; 3 SG<3SG)
dem+subs	–	1 example
dem+subs-NUM	–	1 example (PL)
dem+subs-POSS	–	2 examples (SG<3SG)
dem+adj+subs-POSS	–	1 example (SG<3SG)
dem+subs+subs-POSS	–	1 example (PL<3SG)
num+subs	–	1 example
subs+subs+ptcp+subs-NUM	–	1 example (PL)
subs+pstp+ptcp+subs-NUM	–	1 example (PL)
ppron+subs-POSS	–	1 example (SG<1SG)

Let's turn to the lexically overt DO expressed with pronouns. A declined demonstrative pronoun (*tʲuːt* 'that'; *tʲuːt-ɬəɬ* 'that-PL<3SG') occurs in the sole example where the DO is both lexically and grammatically overt (18) and the DO is a demonstrative. There are no examples in the corpus of lexically and grammatically overt DO with the form of a personal pronoun.

(18) *А* ***тьу*** ***айӄө-нə*** *оказывается* ***тьут-ԓаԓ***
and **that** **young_man-LOC** turns_out **that-PL<3SG**
əйнам *йӑӄəнам* ***аԓ-ԓ-əԓəԓ*** *тьу* *пөп* *мəта*
all home **take-PRS-PL<3SG** that priest some
ӫт-əт
thing-PL
'And that young man, it turns out, takes all of his home, those some kind of things of the priest.' (Schön 2023)

The demonstrative as DO in (18) summarizes the DOs mentioned beforehand in the tale and is also explained at the end of the sentence, in verb-final position. The locAg in this sentence is clearly animate (*тьу*

айқөнə ‘(by) that young man’) and the DO is inanimate (*тьутлалӈ* ‘his things [lit. his many those]’), and as such, it is not higher in animacy.

Concerning the two examples with a DO as a personal pronoun, we have already seen one of them in (6), where both the locAg and the DO are human beings. As they are husband and wife, they are both equally animate. They are assumed to be equally topical as well. The other example in the corpus with locAg in an active sentence and a DO expressed by a personal pronoun (19) is pragmatically slightly different[25], as the DO (*məŋɐt* ‘us’) is indeed higher in the animacy hierarchy (both locAg and DO are human beings, but the former is expressed with a modified noun and the latter with a 1PL.ACC personal pronoun). Furthermore, the locAg in this sentence (19) is expressed by a swear word (*tʲiː jimpəʃi otnə* ‘(by) this poor little soul’, with the human being dehumanized by referring to it with the word *ot* ‘thing, creature’), which designates a man gone mad and posing most probably a menace to the others, behaving normally, and is so clearly lower in animacy than the DO.

(19) *tam* *tʲuː* ***tʲiː*** ***jimpəʃi*** ***ot-nə*** *tam*
then that **this** **unimaginable_thing** **thing-LOC** then
məŋɐt *qotə* *əntə* ***βɛr-t***
1PL.ACC how NEG **do-PRS.3SG**
‘Then this poor little soul will by the end hurt us!’
(Schön 2017e)

Interestingly, all the 26 examples with locAg in active sentences and a lexically or a lexically and grammatically overt DO have an animate locAg, but the DO is slightly more often inanimate than animate (inanimate DO 15/26; 58%; animate DO 11/26, 42%; see also Table 12). So, the DO is not necessarily superior in animacy.

The data of my corpus clearly shows that it is not necessarily the competition between two topics (the subject/locAg and DO/patient) which triggers the locative marking of the subject. There are even examples with no DO at all. My findings confirm the results of Sosa (2017) and Filchenko (2006) presented above that there are diverse possibilities to express the DO and if it is lexically present, it is mostly an identifiable noun phrase.

[25] Grammatically the locAg and DO in the two examples (6) and (19) are similar as they are expressed in both cases with a modified noun in locative and personal pronoun in accusative, respectively.

3.9. Information structure

The most complex question concerning the locAg in Yugan Khanty active sentences, namely its role from the point of view of information structure, is the least fully elaborated one in this paper, and will have to be analyzed in depth in another study, but preliminary results with examples can be shown here already.

The literature states that the locAg is mostly used in topic switch contexts (Filchenko 2006: 47; Sosa 2017: 206; Klumpp – Skribnik 2022: 1024) or in the case of emphasized subjectness (Sosa – Virtanen 2018: 245). My data shows that it finds a wider use than topic switch contexts or emphasis on referents. An example for the first occurrence of an important topic (7c), where a new referent is introduced as brand new information in the third sentence of a tale and is in focus position as a subject of the event-reporting sentence (Lambrecht 1994), was already presented above (7a–c), and an example of topic continuity (9a–c) was also discussed earlier. Even if in (9c) a new referent appears (*Аңкэм* ‘my mother’), I am arguing here against topic switch and for topic continuity, as the main referent is interpreted with the topic definition of Lambrecht. So, Uncle-Demyanka remains the topic of the sentence as we do not get any information about the mother.

There is an even more stunning example for the locAg in Yugan Khanty active sentences than (7c), namely example (20), representing brand new information, as it occurs in the first sentence of a tale, which is without doubt a sentence-focus sentence (Lambrecht 1994).

(20) *Әй* *мәта* *ӆат-нә* ***мәӽ*** ***өвтынә***
one some_kind_of time-LOC **land** **on_the_surface_of**
йӑӈкилӆ-тә ***войӽ-әт***
go_to_do_smth-PTCP.PRS **animal-PL**
саӈки ***ӆэв-нә*** ***ӆәӽӆәӽӆә-тә*** ***тӧвӆ-әӈ*** ***войӽ-әт***
heaven **wave-LOC** **fly-PTCP.PRS** **feather-PROPR** **animal-PL**
Тӧрәм-нә *тӧвә* *ӆӱваты* ***вуӽ-ӆәӆ***
Torem-LOC there 3SG.DAT **summon.PST-PL<3SG**

‘Once upon a time Torem [the highest God] summoned the animals living on the ground and the birds flying in heaven to himself.’ (Schön 2023)

In addition to being the first sentence of the tale, the word order of (20) has to be pointed out. It is the second example in the corpus for the DO (in this case two DOs: *мәҳ өвтынә йӑңқилтә войҳәт* 'animals living on the ground' and *саңки лэвнә лəҳлəҳлəтә тӗвлəң войҳəт* 'birds flying in heaven') preceding the locAg (*Тӗрəмнə* '(by) Torum') and the active verb in objective conjugation (*вуҳлəл* 'he summoned them'; OSV). This emphasizes even more the sentence-focus nature of the first sentence of the tale.

It could be assumed that the three above-mentioned cases (7c), (9a–c), (20) represent a similar context, and the active sentences with the locAgs are event-reporting sentences (Lambrecht 1994) emphasizing the importance of the referent being salient[26] as well.

In the corpus used for this study there are several examples of reoccurring topic. Here I present an example where both the locAg and the DO in the active sentence are reoccurring topics (21) and a second example where only the locAg is returning (22).

(21a) ... *əj tɐːtnə ottə əj βɒːtʃɐ joɣot*
'... one day er, he came to a town.'

(21b) *tʲuː βɒːtʃnə ɐːrjɐtəttəɣ teːmi ottə torəm qɐːn ottə ɛβi tajɐt*
'In that town he looks around: It comes to light er, that the tsar er has a daughter.'

(21c) *ontətnə noməqsəɣət mɐː ottə tʲi torəm qɐːn ɛβi βətəm sɛr*
'He thinks to himself: "I er will now take this tsar's daughter!"'

(21d) *pɐːn ottə*
'And er.'

(21e) *puːt ɒːtəŋ βɒːtʃ ɒːtəŋ oɣt iːmiːn iːkiːnɐ toɣə taŋ*
'To the wife and the husband ogt... from the end of the village, from the end of the town, there he stepped in.'

[26] Interestingly, Filchenko finds in Vasyugan Khanty the locAg in active sentences less salient (Filchenko 2011: 74).

(21f) *pɯːri-t-əɣət* *tʲi* *torəm* ... ***torəm*** ***qɒːn-nə*** *ottə*
ask-PRS-DU<3SG this tsar ... **world** **tsar-LOC** er
tʲi ***εβi-t*** *otəŋ* *ottə* ***mə-t-təɣ*** *əntə*
this **daughter-SG<3SG** whether er **give-PRS-SG<3SG** NEG
'He asks them: "This tsa... tsar er whether this daughter er, would he give her, or not?"' (Schön 2017f)

In (21f) we find the locAg (*torəm qɒːnnə* '(by) the tsar') with an active verb in objective conjugation (*məttəɣ* 'he gives her') and a DO (*tʲi εβit* 'this daughter of his'). The locAg is already introduced in (21b) by the same compound (*torəm qɒːn* 'the tsar') and in (21c) as part of another compositum (*tʲi torəm qɐːn εβi* 'this tsar's daughter'). The DO also appears already in (21b) with a noun (*εβi* 'daughter') and in the beforementioned compound in (21c). They are both equally topical. The use of the locAg in sentence (21f) can reinforce the salient character of this referent. The third topic, the subject in sentences (21a–c) and (21e) is the main protagonist of the tale (*ɐːtiɬ qɒːtəŋ mɒːnʲtʲ qoː* 'The fairytale hero living alone') and is expressed in all sentences by the 3SG verbal ending and remains topical throughout.

Example (22g) contains a locAg (*tʲuː qoːnə* '(by) that man') as a reoccurring topic, but the DO (*kɯ̈rət* 'his leg') is in focus position, and, accordingly, the verb shows subject agreement, not subject-object agreement. Furthermore, the DO has a possessive suffix, which can function as a focus marker as well (Janda 2019). Intriguingly, there is a self-repair in the narrative (22g), where the storyteller mixes up the protagonists, but both the word said first (*tʲuː neːnə* '(by) that woman') and its correction (*tʲuː qoːnə* '(by) that man') are locAgs.

(22a) *pɐːn tʲuː qoː tot otɐt*
'And that man lies there.'
(22b) *pɐːn kɯ̈rət pɐːn ubral*
'And then he removed his legs.'
(22c) *kɯ̈rət iːtə tuːβtəɣ*
'He removed his legs.'
(22d) *tʲuː iːmi mən*
'And that woman left.'
(22e) *pɐːn tom qɒːtɐ jaŋqətəɣ pɐːnə ottə tiːtot βij*
'And she went to the neighboring house and er took food.'
(22f) *pɐːn piryinɐm lɯːpɣəttətəɣ*
'And she limped back.'

(22g) *pirɣinɐm* *lɯːpɣəttətəɣ* *pɐːn* *tʲuː* *neː-nə* ...
back limp.PST.3SG and that woman-LOC …
tʲuː ***qoː-nə*** *pɐːnə* ***kɯr-ət*** ***iːtə*** ***nɯrtəɣ***
that **man-LOC** then **leg-SG<3SG** **PFV** **stretch.PST.3SG**

‘She limped back and that woman... then that man had stretched his legs.’ (Schön 2016c)

The locAg in (22g) is one of the main human protagonists of the legend, where two men argue and one goes to spend the night in the cemetery. In (22a) he is depicted with a demonstrative and a noun (*tʲuː qoː* ‘that man’), in (22b) and (22c) in the 3SG verbal ending. He reoccurs in (22g) as a locAg in an active sentence, showing the main event of the plot: he stretches out his leg across the path where the dead woman will return to her grave and fall. Even though the DO (*kɯrət* ‘his leg’) of the locAg sentence with an active verb was already introduced in (22c) and (22d), it is in (22g) in focus position. At first sight, the verb (*iːtə nɯrtəɣ* ‘he stretched out’) in (22g) could be in objective conjugation as the *-təɣ* ending is identical to the SG<3SG verbal suffix, but the DEWOS (1015) clearly states that *nɯrt-* is the verbal stem, so the verbal form *nɯrtəɣ* has to be in subjective conjugation, past tense. The DO’s (*kɯrət* ‘his leg’) importance can be emphasized by the fact that it is the connecting point between the worlds of the living and the dead (Dudeck et al. 2021), so its focal position is not only grammatically justified. Interestingly, in one version of the same plot from a different storyteller, the protagonist of this main event is also expressed by a locAg in an active sentence (14), but the leg (also DO) is not in focus position. It rather seems to be a new topic, as it is only introduced in that sentence of the legend (14) and nowhere before (Kayukova – Schön 2020a).

The few data analyzed in terms of information structure and presented here shows that the locAg in Yugan Khanty active sentences appears if the subject should be highlighted as salient or if it is emphasized for any reason. The surprising examples with the locAg in focus position suggest that the locAg finds wider use than assumed based on the data till today. To what extent saliency and topicality play a role in the use of the locAg in Yugan Khanty active sentences requires further, deeper research, but information structure motivated factors seem to determine the choice of the locAg in Yugan Khanty active sentences.

3.10 Animacy and summarized presentation of the data

Last, I briefly want to present all examples containing a locAg in Yugan Khanty active sentences of the corpus used for this study, even if they are not cited as examples in this paper (see Table 12). New factors like animacy and the human character of the locAg and the DO complete the analyses. This is relevant for reasons of topicality and hierarchy (compare Klumpp – Skribnik 2022: 1024 and sections **3.7.**, **3.8.**).

Table 12: Summary of the data used in this study

	Informant	locAg			Lexically overt DO			Conjugation	Word order
		POS[27] +LOC	Animate	Human	POS	Animate	Human		
1/(3)[28]	TMJ	ppron	yes	no	–	–	–	obj.	SV
2/(4?)	AIK	ppron	yes	yes	MN	no	no	obj.	OSV
3/(5)	TMK	ppron	yes	no	–	–	–	subj.	SV
4/(6)	VIU	MN	yes	yes	ppron	yes	yes	subj.	SOV
5/(7c)	AIK	N	yes	no	–	–	–	subj.	SV
6/(8)	AIK	N	yes	no	–	–	–	subj.	SV
7/(9c)	TMJ	MN	yes	yes	–	–	–	subj.	SV
8/(10)	ENK	ppron	yes	no	MN	no	no	obj.	SOV
9/(11)	MDA	ppron	yes	yes	–	–	–	obj.	VS
10/(12)	TMJ	MN	yes	yes	MN	no	no	obj.	OVS
11/(13)	ENK	ppron	yes	no	MN-POSS	no	no	obj.	OVS
12/(14)	ANB	MN	yes	yes	N-POSS	no	no	obj.	OVS
13/(15b?)	ENK	ppron	yes	yes	N	no	no	subj.	SOV
14/(17)	VIU	N-POSS	yes	yes	–	–	–	subj.	SV
15/(18)	LDK	MN	yes	yes	dem-POSS	no	no	obj.	SOV
16/(19)	VIU	MN	yes	yes	ppron	yes	yes	subj.	SOV
17/(20)	LNK	PN	yes	no	2 MN	yes	no	obj.	OSV
18/(21f)	VIU	N	yes	yes	MN-POSS	yes	yes	obj.	SOV
19/(22g)	ENK	MN	yes	yes	N-POSS	no	no	subj.	SOV
20/–	VIU	MN	yes	yes	N-POSS	yes	yes	subj.	OVS
21/–	AIK	MN	yes	yes	N	yes	yes	obj.	SOV
22/–	AIK	N	yes	yes	–	–	–	obj.	SV

[27] MN = modified common or proper noun, N = non-modified common noun, PN = proper noun.

[28] The first number is a serial number of the 41 examples from the corpus used for this study, the second number is the number of the example if cited in this paper.

	Informant	locAg			Lexically overt DO			Conjugation	Word order
		POS[29] +LOC	Animate	Human	POS	Animate	Human		
22/–[30]	AIK	N	yes	yes	–	–	–	obj.	SV
23/–	AIK	ppron	yes	no	–	–	–	obj.	SV
24/–	AIK	MN	yes	no	MN	no	no	obj.	OVS
25/–	VIU	N	yes	yes	MN-POSS	yes	yes	obj.	SOV
26/–	ENK	N	yes	yes	N-POSS	yes	yes	obj.	SOV
27/–	VIU	N	yes	yes	N-POSS	yes	yes	obj.	SOV
28/–	ANA	MN	yes	yes	MN-POSS	yes	yes	obj.	SOV
29/–	ANA	MN-POSS	yes	no	–	–	–	obj.	SV
30/–	LDK	ppron	yes	yes	N-POSS	no	no	obj.	SOV
31/–	VIU	MN	yes	yes	N	no	no	subj.	SOV
32/–	VIU	MN	yes	yes	–	–	–	subj.	SV
33/–	TMJ	MN	yes	yes	–	–	–	obj..	SV
34/–	TMJ	MN	yes	yes	–	–	–	obj.	SV
35/–	VIU	interro	yes	yes	–	–	–	subj.	SV
36/–	TAK	N	yes	no	MN-POSS	no	no	obj.	SOV
37/–	ANA	MN-POSS	yes	yes	–	–	–	subj.	SV
38/–	MDA	MN	yes	no	N-POSS	no	no	obj.	SOV
39/–	MDA	MN	yes	no	N-POSS	no	no	obj.	SOV
40/–	MDA	MN	yes	no	N-POSS	no	no	obj.	SOV
41/–	AIK	MN	yes	yes	–	–	–	obj.	SV

As we have already seen, the locAg can be expressed by a personal pronoun, an interrogative pronoun, a proper noun, or a common noun (with or without modifiers and with or without a possessive marker). In all cases the locAg in active sentences is animate, but not in all cases human (14/41; 34%). Due to the genre of tales and legends, it can be a deity (Torum, the highest Khanty God; 17/(20)), an animal protagonist (like a rabbit; 1/(3), 24/–) or a humanized, talking object (e.g., a needle; 36/–). The DO is not obligatory, but if it is lexically overt (25/41; 61%), it can be expressed by a personal pronoun, a demonstrative, or a common noun (with or without modifiers and with or without a possessive marker).

[29] MN = modified common or proper noun, N = non-modified common noun, PN = proper noun.

[30] The first number is a serial number of the 41 examples from the corpus used for this study, the second number is the number of the example if cited in this paper.

Interestingly, there are more DOs expressed with possessive marked non-modified nouns (9/26; 35%) than with bare nouns (3/26; 12%), but an almost equal number of modified nouns with or without possessive suffixes (5/26; 19% and 6/26; 23%). The lexically overt DOs can be inanimate (15/26; 58%; 8/(10), 10/(12)) as well as animate (11/26; 42%; 4/(6), 18/(21f)). There are only two cases, in which the DO is animate and not human (animals and birds; 17/(20)). In all other cases, if the DO is animate, then it is human (18/(21f)) and if it is inanimate, then it is non-human (8/(10)). These examples show that the factor of animacy of the subject or the direct object is not decisive in the use of the locAg in Yugan Khanty active sentences.

4. Some possible explanations of the origin of locAg in Surgut Khanty active sentences

The question of the origin of the locAg in Surgut Khanty active sentences remains open even after different analyses by several researchers. I will propose some highly hypothetical explanations.

As mentioned in the Introduction (**1.**), I have kept the name locative-marked agent instead of locative-marked subject to highlight its connection with active (1), (3) and passive sentences (2), as in the latter locAgs can appear besides the grammatical subjects. In my corpus used for this paper, there are over 400 examples with a locAg in passive sentences. Compared to the 41 locAgs in active sentences presented above, it would be unreasonable to say that the locAg in active sentences is taking the place of the locAg in passive sentences. But looking at all the locAgs in the corpus used for this study, there might be some connection. In example (2) we have a neutral, typical example for the grammatical subject (‘the moon and the sun’), a locAg (‘by his daughter-in-law’) and a transitive verb in passive voice (‘hide’). In examples (3–15b?) and (17–22g) we see different realizations of the locAg in active sentences. Example (23) demonstrates the locAg with a present participle inflected with a personal ending in 1DU (*joɣttɐmn* ‘we two coming’), but it is clear at first sight that the locAg (*βoqinə* ‘by the fox’) does not show any agreement with the participle. This happens twice, as the self-repair here is only semantic (the two rabbits are followed by the fox), not grammatical.

(23) *oːs* *tʃeːβər-nə* *joɣt-t-ɐmn* *kɐː*
more rabbit-LOC come-PTCP.PRS-1DU when
βoqi-nə *joɣt-t-ɐmn* *kɐː*
fox-LOC come- PTCP.PRS-1DU when
noq *ɬiː-ɬ-ojmən* *βiːɬe*
PFV eat-PRS-PASS.1DU in_fact

'Otherwise, when we are outrun by the rabbit, when we are outrun by the fox, we will be eaten up after all.'
(Kayukova – Schön 2020b)

In (24) we find another example with locAg (*tʲuː iːminə* '(by) that old woman') and past participle (*pɐːrtmɐɬ* 'she ordered'), but here the participle shows agreement with the locAg.

(24) *pɐːnə* *tʲuː* *tʃeːmotɐn-əɬ* *tʲuː* *iːmi-nə*
and that suitcase-SG<3SG that old_woman-LOC
pɐːrt-m-ɐɬ *sɒːɣət* *pɐːnə* *quːtʲŋ-əɬ-ɐ*
order-PTCP.PST-3SG as then beside-3SG-LAT
uːmət-təɣ
put_down.PST-SG<3SG

'Following the old woman's instructions, she laid the suitcase next to her. [Lit. And her suitcase, as that old woman had ordered, she put it down next to her.]' (Csepregi 1998d)

Intriguing in (24) is the person marking at the end of the participle (*-ɐɬ* '-3SG'), because as the referent is lexically overt, the sentence would be grammatical and comprehensible without it, too. If this sentence (*tʲuː iːminə pɐːrtmɐɬ sɒːɣət* 'as that old woman had ordered') were clearly a passive one, then there would be no personal marking on the participle (e.g., with the locative suffix staying at the agent *tʲuː iːminə pɐːrtəm sɒːɣət* or even without the locative suffix *tʲuː iːmi pɐːrtəm sɒːɣət*). Personal marking on participles is only obligatory if the agent is expressed by a personal pronoun or is not lexically overt – both in active constructions (Csepregi 1978, 1979; and personal account in 2022).

While in (23) the locAg appears besides the subject as a second referent and encodes the source of fear, in (24) the locAg depicts a different, prominent referent and the subject of the participle.

In my opinion, the following series of changes is possible:

A) the locAg is available in passive sentences (2) →
B) then the locAg appears with participles as second referent (23) →
C) after this, the locAg appears with participles as (prominent) referent and the participle shows agreement with it (24) →
D) finally, the locAg appears in active sentences and shows agreement with the active verb (3–15b?) and (17–22g).

Clearly, this hypothesis requires further research.

Another possible explanation came up in the discussion of my talk presenting the locAg in active sentences in February 2022, by Lena Borise. The idea arose because of the appearance of the locAg as a 1SG.LOC personal pronoun (see footnote 16) and its homonymy with the 1SG.EMPH pronoun (see Table 13). As locAgs seem to have emphatic nuances or topical meanings too, this homonymy might not be a coincidence. The supposed first appearance of the locAg in Paasonen's collection (16) is also in the 1SG.LOC/1SG.EMPH form, which could also confirm the connection.

Table 13: Possible meaning change of *mɐ̆ːnə*

mɐ̆ː mɐ̆ːnə	>	*mɐ̆ːnə*	>	*mɐ̆ːnə*
1SG.EMPH	>	1SG.EMPH	>	1SG.LOC
'I myself'	>	'myself'	>	'by me'

There are a long and a short form for the emphatic personal pronoun (*mɐ̆ː mɐ̆ːnə* and *mɐ̆ːnə*), both used in Surgut Khanty narratives. The ending *-nə* of the 1SG emphatic personal pronoun is the same as the ending of the locative suffix. So, the short form of the 1SG emphatic personal pronoun could have been reanalyzed as a 1SG personal pronoun in locative (1SG.LOC), and this locative marking could have spread to other subjects, which would thereby become locAgs indicating the emphatic meaning. Whether this homonymy is the source of the change itself, or one additional factor of the emergence of locAgs, or indeed just pure chance, is not clear at this point of time. Its significance requires further research.

5. Summary

To sum up all the observations presented in **3.1–3.10.**, I will recapitulate the factors surveyed in order to show which of them turned out to be relevant or irrelevant, and I will highlight further similarities with and differences from the results of other researchers dealing with the topic of

the locAg in active sentences in Surgut or Vakh-Vasyugan Khanty, namely Sosa (2017), Filchenko (2006, 2011), and Csepregi (2020, 2021).

An idiolectal motivation of the appearance of the locAg in Yugan Khanty active sentences is clearly disproved, as there are 11 speakers in the corpus making use of it. The 11 storytellers differ in all possible sociolectal factors (**3.1.**). As they all come from different parts of the Yugan Khanty language area (from the upper and lower streams of the rivers Bolshoy Yugan and Malyy Yugan through the river Salym to the Demyanka), subdialectal reasons for the use of the locAg in active sentences is excluded, too, as well as other (sub)dialectal reasons, since the phenomenon appears in all other Surgut Khanty subdialects (**3.2.**). The influence of ritual language is also highly doubtful as an explanation, inasmuch as the locAg in active sentences appears extremely rarely in the language of Surgut Khanty ritual genres. A contact phenomenon from Vakh-Vasyugan ritual language can also be excluded since the two dialects are not mutually intelligible and most of the storytellers of my corpus do not understand the Vakh-Vasyugan ritual language (**3.3.**).

No common characteristics of the verbs used in active and locAg constructions could be found. Concerning verbal semantics, 30 different verbs with 25 different meanings appear in the examples, with no semantic restriction that could be determined (**3.4.**). Regarding transitivity, both transitive and intransitive verbs are present with the locAg in active sentences in the corpus, but transitive verbs clearly dominate (38/41; 93%). One intransitive motion verb (‘come’) shows up three times in examples (7c), (8), (9c). The verbs manifest subject agreement (subjective conjugation; 13/41; 32%), as well as subject-object agreement (objective conjugation; 28/41; 68%). So, neither transitivity, nor conjugation is a determining factor in the choice of a locAg over an unmarked subject (**3.5.**).

The word order of the sentences containing a locAg and an active verb mainly coincides with the neutral Khanty SOV word order (33/41; 80%). The remaining 20% of the examples (8/41) are analyzed in detail in section **3.6.** Some examples possibly show Russian influence, some are an afterthought, and in some cases emphasis appears. Interestingly, all 8 examples with non-SOV word order contain a verb with subject-object agreement.

The internal structure of the subject in Yugan Khanty as a locAg in active sentences is always overt and mostly already accessible or identifiable by its grammatical form – like a personal pronoun (9/41; 22%; (3),

(5)); a proper noun (2/41; 5%; (9c), (20)); or common noun with a demonstrative (20/41; 49%; (14), (19)). Surprisingly, and contrary to the observation of the literature, non-modified nouns in the locative also appear as locAgs in active sentences (8/41; 20%; (7c), (8)). That means that there is no part-of-speech restriction on the use of the locAg in Yugan Khanty active sentences (**3.7.**). Problematic are the examples containing a locAg expressed with a locative personal pronoun in 1SG (*mɐ̆ːnə*), because of its homonymy with the 1SG emphatic personal pronoun (*mɐːnə*). All examples containing *mɐːnə* are at least ambiguous. Fortunately, this concerns only two examples from among those used for this study ((4?), (15b?); **3.7.**).

The internal structure of the direct object in Yugan Khanty as a locAg in active sentences shows that the DO is mostly expressed with some kind of a noun (with or without modifiers and with or without a possessive marker) (23/26; 88%), and only in a few cases with pronouns (3/26; 12%). The last point is surprising, because it contradicts other statements, and shows that the part-of-speech affiliation of the DO does not play any role in the use of the locAg. Even more amazing is the fact that not all examples in my corpus have to have a DO of any kind (lexically or grammatically overt), and the locAg in the active sentence is still present. So, the assumption that locAg in Khanty active sentences appears in the case of two concurring topics to highlight which is higher in animacy is not in all cases true (**3.8.**, **3.10.**).

Even though information structure is the least elaborated motivation in this study, it seems to be the only relevant factor in the choice of the locAg in Yugan Khanty active sentences. In addition to topic switch contexts and emphasized subjectness, its use to introduce brand new information (7c), (20) or its appearance in a topic continuity context (9c), as well as reoccurring topic (21f), (22g) has been pointed out.

The following conclusions can be drawn from the data. Most of the factors examined play little or no role, including: idiolectal motivation (**3.1.**), dialectal or subdialectal motivation (**3.2.**), influence from ritual language (**3.3.**), verbal semantics (**3.4.**), transitivity or conjugation (**3.5.**), word order or word order changes (**3.6.**), internal structure of the subject (**3.7.**) or of the direct object (**3.8.**), as well as their animacy (**3.10.**). The only relevant factor analyzed is information structure (**3.9.**). The data used in this study clearly reflects that encoding information structure, emphasis on the referent, and saliency play a role in the choice of locAg in Yugan Khanty active sentences.

The most obvious difference between the research so far in Surgut Khanty and my analysis above (**3.1**–**3.10.**) is the number and certainty of the data used (see Table 14). The corpus used in this study containing 41 examples of the locAg in active sentences forms the Surgut Khanty corpus with the highest amount of data for this linguistic phenomenon. Compared to this, Csepregi presents one example emerging in a Tromagan Khanty ritual song in the collection of Karjalainen (Csepregi 2020: 45–46) and Sosa works with nine examples (2017: 182) – one collected by herself and eight found in Csepregi's collections, namely Csepregi (1998a) and (2011) (see Sosa 2017: 70–72, 188–191).

Table 14: Summary of data used for locAg in Surgut Khanty active sentences

	Number of examples
Sosa 2017	9
Csepregi 2020	1
Schön 2022	41

The researchers mentioned in the first paragraph of this section state unanimously that the locAg in active sentences is always animate. As shown in Table 12, my data supports this to 100% and proves that the locAg in Yugan Khanty active sentences does not necessarily have to be human, just animate. In the data used by Sosa and Filchenko, the locAg is mostly expressed by personal pronouns (Sosa 2017, Filchenko 2006), which remains in opposition to my findings (see section **3.7.** and Table 12).

Sosa finds that the object in an active sentence with locAg is mostly grammatically overt. Filchenko (2006) states that the object, if present, is an identifiable noun phrase, which is confirmed by my data, except that I have found examples with a personal pronoun in the accusative case as well for the direct object like Filchenko (2011) did (see section **3.8.** and Table 12, as well as (6), (19)).

Csepregi (2020, 2021) observes that the active verb with a locAg can be in subjective and objective conjugation as well. This is also what my results show (see section **3.5.** and Table 12).

Csepregi (2021) suggests that the locAg in active sentences can be used as a stylistic device to alternate between different constructions and to avoid repetition, but this is not attested in my data (see section **3.3.**).

The finding of Filchenko that the locAg in active sentences in Vasyugan Khanty is in connection with an unintended action or with "the degree of the participants' agentivity, control and volition" (Filchenko

2006: 47; as well as Filchenko 2011: 63, 71–73), is not entirely supported by my data and needs further research.

Concerning information structure, most of the researchers working on the locAg agree that the locAg appears in topic switch contexts (Filchenko 2006, Sosa 2017, Sosa – Virtanen 2018,) or in cases of topic continuity (Sosa 2017). Csepregi (2020, 2021) has found additionally that the locAg can be a reoccurring agent and can appear in consecutive sentences. As demonstrated in section **3.9.**, my preliminary analysis confirms all these claims as well. Sosa (2017) states further that the locAg is the topic of the sentence; according to my data it can be in focus position as well (see section **3.9.**).

One important similarity with Sosa (2017) is her finding, contrary to Kulonen (1989) and Csepregi (2021), that the locAg in active sentences occurs not only with transitive verbs, but with intransitive verbs as well (see sections **3.6.** and **3.7.**).

In this data-driven pilot study, the locAg in Yugan Khanty sentences was presented from several different points of view. It is quite clear which factors do *not* play a role in its use. Some perspectives of it, however, like the question of topicality, saliency, or its role in information structure, to name a few, require further research.

References

Asztalos, Erika – Gugán, Katalin – Mus, Nikolett (2017), Uráli VX szórend: nyenyec, hanti és udmurt mondatszerkezeti változatok. In: É. Kiss, Katalin – Hegedűs, Attila – Pintér, Lilla (eds), Nyelvelmélet és diakrónia 3. PPKE BTK Elméleti Nyelvészeti Tanszék – Magyar Nyelvészeti Tanszék, Budapest – Piliscsaba. 30–62.

Balalaeva, O. E. – Surlomkina, E. P. – Wiget, E. O. [Балалаева, О. Э. – Сурломкина, Е. П. – Уигэт, Э. О.] (2021), Голоса Югана. Сборник фольклора Йавэн-Йах. Печатный мир, Сургут.

Balandin, A. N. [Баландин, А. Н.] (1948), Падежи субъекта на службе номинативной и эргативной конструкции глагольного предложения в ваховском диалекте хантыйского языка. Советское Финно-угроведение 1. Ученые записки ЛГУ 105. Серия востоковедческих наук, вып. 2: 301–315.

Csepregi, Márta (1978), Keleti osztják igeneves szerkezetek. Nyelvtudományi Közlemények 80: 31–53.

Csepregi, Márta (1979), Keleti osztják igeneves szerkezetek II–III. Nyelvtudományi Közlemények 81: 273–300.

Csepregi, Márta (1998a), Szurguti osztják chrestomathia. Studia Uralo-Altaica. Supplementum 6. JATE Finnugor Tanszék, Szeged.
Csepregi, Márta (1998b), OUDB Surgut Khanty Corpus. Text ID 728. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=728 [Last access: 30-09-2022]
Csepregi, Márta (1998c), OUDB Surgut Khanty Corpus. Text ID 730. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=730 [Last access: 30-09-2022]
Csepregi, Márta (1998d), OUDB Surgut Khanty Corpus. Text ID 735. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=735 [Last access: 30-09-2022]
Csepregi, Márta (2011), Szurguti hanti folklór szövegek. Budapesti Finnugor Füzetek 22. ELTE BBTK Finnugor Tanszék – Numi Tórem Finnugor Alapítvány, Budapest.
Csepregi, Márta [Чепреги, Марта] (2017), Сургутский диалект хантыйского языка. Обско-Угорский Институт Прикладных Исследователя и Разработок, Ханты-Мансийск.
Csepregi, Márta (2020), K. F. Karjalainen keleti hanti szöveggyűjtése 1899–1901. Akadémiai doktori értekezés, Budapest.
Csepregi, Márta (2021), Jelölt ágens aktív ragozású igék mellett a hantiban. Segítenek-e tisztázni az ún. ergatív típus kérdését Karjalainen vaszjugani szövegei. Előadás, 2021. március 2.
DEWOS = Steinitz, Wolfgang et al. (1966–1993), Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache. Akademie-Verlag, Berlin.
Dudeck et al. = Dudeck, Stephan – Schön, Zsófia – Vagt, Christian – Pesikova, Agrafena – Kechimov, Yosif – Usanov, Vasilii – Vella, Yuri (2021), Troubling visits and uncanny encounters – Indigenous concepts of other than humans and their homes. In: TRANSLOCAL. Culturas Contemporâneas Locais e Urbanas, n.º5 – Espaços (Des)Habitados | (Un)Inhabited Spaces, Funchal: UMa-CIERL/CMF/IA. translocal.cm-funchal.pt/2019/05/02/revista05/. https://translocal.cm-funchal.pt/wp-content/uploads/2019/05/Troubling-Visits-and-Uncanny-Encounters_5TRANSLOCAL.pdf. [Last access: 30-09-2022]
Filchenko, Andrey (2006), The Eastern Khanty locative-agent constructions. A functional discourse-pragmatic perspective. In: Lyngfelt, Benjamin – Solstad, Torgrim (eds), Demoting the agent. Passive, middle and other voice phenomena. Linguistik Aktuell/Linguistics Today 98. John Benjamins, Amsterdam. 47–82.
Filchenko, Andrey Y. (2011), Parenthetical agent-demoting constructions in Eastern Khanty: Discourse salience vis-à-vis referring expressions. In: Chiarcos, Christian – Berry, Claus – Grabski, Michael (eds), Salience: Multidisciplinary perspectives on its function in discourse. De Gruyter Mouton, Berlin – New York. 57–80.

Gugán, Katalin – Sipos, Mária (2017), Ige mögötti mondatrészek régi hanti szövegekben. In: Forgács, Tamás – Németh, Miklós – Sinkovics, Balázs (eds), A nyelvtörténeti kutatások újabb eredményei IX. SZTE Magyar Nyelvészeti Tanszék, Szeged. 77–94.

Gulya, János (1970), Aktiv, Ergativ und Passiv im Vach-Ostjakischen. In: Symposion über Syntax der uralischen Sprachen. 15-18. Juli 1969 in Reinhausen bei Göttingen. Abhandlungen der Akademie der Wissenschaften in Göttingen. 80–83.

Gulya, János (1982), Das Ergativproblem in den finnisch-ugrischen Sprachen. Ural-Altaische Jahrbücher. Neue Folge 1: 19–34.

Gulya, János (1994), A mondatszerkezetek rendszere a vahi osztjákban. Studia Uralo-Altaica Supplementum 4. JATE, Szeged.

F. Gulyás, Nikolett (2018), Ditranzitív szerkezetek a keleti hantiban: a *mə-* 'ad' ige. In: Kenesei István (ed.), Általános Nyelvészeti Tanulmányok XXX. Uralisztikai tanulmányok. Akadémiai Kiadó, Budapest. 71–95.

Havas, Ferenc (2003), A tárgy tárgyában. In: Oszkó, Beatrix – Sipos, Mária (eds), Uráli tárgyaló. Budapesti Uráli Műhely III. MTA Nyelvtudományi Intézet, Budapest. 7–44.

Havas, Ferenc (2011), A prenominativitás elmélet és alkalmazása az uráli nyelvekre. Akadémiai doktori értekezés, Budapest.

Honti, László (1971), A cselekvő (logikai) alany alaktana az obi-ugor nyelvekben. Nyelvtudományi Közlemények 73: 430–440.

Honti, László (1984), Chrestomathia Ostiacica. Tankönyvkiadó, Budapest.

Janda, Gwen (2019), Funktionen von Possessivsuffixen in den ugrischen Sprachen. Modern Academic Publishing, Köln.

Kayukova, L. N. – Schön, Zsófia (2016a), OUDB Yugan Khanty (2010–) Corpus. Text ID 1518. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1518 [Last access: 30-09-2022]

Kayukova, L. N. – Schön, Zsófia (2016b), OUDB Yugan Khanty (2010–) Corpus. Text ID 1515. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1515 [Last access: 30-09-2022]

Kayukova, L. N. – Schön, Zsófia (2017), OUDB Yugan Khanty (2010–) Corpus. Text ID 1615. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1615 [Last access: 30-09-2022]

Kayukova, L. N. – Schön, Zsófia (2019), OUDB Yugan Khanty (2010–) Corpus. Text ID 1648. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1648 [Last access: 30-09-2022]

Kayukova, L. N. – Schön, Zsófia (2020a), OUDB Yugan Khanty (2010–) Corpus. Text ID 1658. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1658 [Last access: 30-09-2022]

Kayukova, L. N. – Schön, Zsófia (2020b), OUDB Yugan Khanty (2010–) Corpus. Text ID 1660. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1660 [Last access: 30-09-2022]

Klumpp, Gerson – Skribnik, Elena (2022), Information structuring. In: Bakró-Nagy, Marianne – Laakso, Johanna – Skribnik, Elena (eds), The Oxford guide to the Uralic languages. Oxford University Press, Oxford. 1018–1036.
Koshkareva, N. B. (ed.) [Кошкарёва, Н. Б. (сост.)] (2006), Детские сказки варёганских ханты (на хантыйском и русском языках). Полиграфист, Ханты-Мансийск.
Kulonen, Ulla-Maija (1989), The passive in Ob-Ugrian. Mémoires de la Société Finno-Ougrienne 203. Suomalais-Ugrilainen Seura, Helsinki.
Kulonen, Ulla-Maija (1991), Über die ergativischen Konstruktionen im Ostjakischen. Journal de la Société Finno-Ougrienne 83: 181–201.
Lambrecht, Knud (1994), Information structure and sentence form. Cambridge Studies in Linguistics 71. Cambridge University Press, Cambridge.
Nikolaeva, Irina (1999), Object agreement, grammatical relations, and information structure. Studies in Language 23: 341–386.
Paasonen, Heikki – Vértes, Edith (2001), OUDB Yugan Khanty (1901) Corpus. Text ID 1316. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1316 [Last access: 30-09-2022]
Pesikova, A. S. – Volkova, A. N. [Песикова А. С. – Волкова А. Н.] (2013), Сказки, рассказы с реки Лямы. Фольклорный сборник на языке сургуртских ханты. Юграфика, Ханты-Мансийск.
Ruttkay-Miklián, Eszter (2003), Az alany és a tárgy jelölése a hanti nyelvben. In: Oszkó, Beatrix – Sipos, Mária (eds), Uráli tárgyaló. Budapesti Uráli Műhely III. MTA Nyelvtudományi Intézet, Budapest. 114–137.
Schön, Zsófia (2015a), Yugan Khanty Personal Archive Schön ID087.
Schön, Zsófia (2015b), Yugan Khanty Personal Archive Schön ID091.
Schön, Zsófia (2016a), OUDB Yugan Khanty (2010–) Corpus. Text ID 1479. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1479 [Last access: 30-09-2022]
Schön, Zsófia (2016b), OUDB Yugan Khanty (2010–) Corpus. Text ID 1514. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1514 [Last access: 30-09-2022]
Schön, Zsófia (2016c), OUDB Yugan Khanty (2010–) Corpus. Text ID 1503. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1503 [Last access: 30-09-2022]
Schön, Zsófia (2016–2017), Tromagan Khanty Personal Archive Schön.
Schön, Zsófia (2017a), OUDB Yugan Khanty (2010–) Corpus. Text ID 1596. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1596 [Last access: 30-09-2022]
Schön, Zsófia (2017b), OUDB Yugan Khanty (2010–) Corpus. Text ID 1617. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1617 [Last access: 30-09-2022]

Schön, Zsófia (2017c), OUDB Yugan Khanty (2010–) Corpus. Text ID 1621. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1621 [Last access: 30-09-2022]
Schön, Zsófia (2017d), OUDB Yugan Khanty (2010–) Corpus. Text ID 1635. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1635 [Last access: 30-09-2022]
Schön, Zsófia (2017e), OUDB Yugan Khanty (2010–) Corpus. Text ID 1627. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1627 [Last access: 30-09-2022]
Schön, Zsófia (2017f) OUDB Yugan Khanty (2010–) Corpus. Text ID 1609. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1609 [Last access: 30-09-2022]
Schön, Zsófia (2019a), OUDB Yugan Khanty (2010–) Corpus. Text ID 1646. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1646 [Last access: 30-09-2022]
Schön, Zsófia (2019b), OUDB Yugan Khanty (2010–) Corpus. Text ID 1644. Schön, Zsófia (ed.), https://www.oudb.gwi.uni-muenchen.de/?cit=1644 [Last access: 30-09-2022]
Schön, Zsófia (2020), (S)he lives with that legend, with that fairytale till today. In: Ёмас сымыӈ нэ̄квe вō̄ртур э̄тпост самын патум – Scripta miscellanea in honorem Ulla-Maija Forsberg. Mémoires de la Société Finno-Ougrienne 275: 348–368.
Schön, Zsófia (2023), Märchen und Legenden vom Großen und Kleinen Jugan. Kulturstiftung Sibirien, Fürstenberg. [Forthcoming.]
Schön, Zsófia – Gugán, Katalin (2022), East Khanty. In: Bakró-Nagy, Marianne – Laakso, Johanna – Skribnik, Elena (eds), The Oxford guide to the Uralic languages. Oxford University Press, Oxford. 608–635.
Sigurðsson, Halldór Ármann (2007), Icelandic non-nominative subjects. In: Bhaskararao, Peri – Subbarao, K. V. (eds), Non-nominative subjects. Vol. 2. John Benjamins, Amsterdam. 137–159.
Sosa, Sachiko (2017), Functions of morphosyntactic alternations, and information flow in Surgut Khanty discourse. University of Helsinki Faculty of Arts, Helsinki.
Sosa, Sachiko – Virtanen, Susanna (2018), Remarks on areal linguistics in the information structure of the Ob-Ugrian languages. Folia Uralica Debreceniensia 25: 233–260.
SVH = Tereshkin, N. I. [Терешкин, Н. И.] (1981), Словарь восточно-хантыйских диалектов. Наука, Ленинград.