# The influence of internal possessors on argument coding in Northern Khanty



**Nikita A. Muravyev**
University of Hamburg, Hamburg, Germany; nikita.muraviev@gmail.com

**Abstract**: The study examines contexts in Northern Khanty where the presence of internal possessors in the noun phrases of the core arguments influences their morphosyntactic coding. Based on the corpus data from the Kazym dialect of Khanty, I investigate contexts in which the presence of a third-person topical possessor in the agent or patient NP leads to agent demotion and passivization. The article provides an analysis of these contexts in terms of the interaction between discourse prominence and processing efficiency. It is demonstrated that passive voice is used as an ambiguity-resolving strategy when several unmarked participants are present in the clause, or when referential conflict may arise due to unclear reference of an anaphoric possessive marker. This strategy applies whenever the agent is less prominent than either the patient or the possessor of the patient, or whenever both possessors of the agent and the patient are present, and both core arguments are low on animacy, agentivity, and topicality hierarchies, in which case passivization is due not to their relative prominence but rather to considerations of processing efficiency.

**Keywords**: argument structure, Northern Khanty, passive, possession, Uralic, voice
**Acknowledgements**: The research was conducted as part of the project "Verb Agreement and Voice in Northern Khanty Discourse", supported by the Alexander von Humboldt Foundation. I am grateful to my colleagues and collaborators, Stiopa Mikhailov and Vsevolod Masliukov, for discussing my data from the perspective of hierarchical alignment and obviation. I also thank the volume editors and the anonymous reviewer who helped improve the quality of this paper. Any remaining mistakes are my own.
**For citation**: Muravyev N. A. The influence of internal possessors on argument coding in Northern Khanty. *Voprosy Jazykoznanija*, 2024, 5: 65–79.
**DOI**: 10.31857/0373-658X.2024.5.65-79

# Влияние внутреннего посессора на кодирование аргументов в севернохантыйском языке

**Никита Алексеевич Муравьев**
Университет Гамбурга, Гамбург, Германия; nikita.muraviev@gmail.com

**Аннотация**: В исследовании рассматриваются контексты в севернохантыйском языке, в которых наличие внутреннего посессора в именных группах ядерных участников влияет на кодирование этих актантов. На основе корпусных данных казымского диалекта хантыйского рассматриваются контексты, в которых присутствие топикального посессора третьего лица в именной группе агенса или пациенса приводит к понижению агенса и пассивизации. В статье проводится анализ таких контекстов с точки зрения взаимодействия дискурсивной значимости и эффективности обработки речи. Установлено, что пассивный залог используется как стратегия разрешения неоднозначности, когда в предложении присутствуют несколько немаркированных участников, а также когда может возникнуть референциальный конфликт из-за неоднозначной референции анафорического посессивного маркера. Эта стратегия применяется, когда агенс менее дискурсивно значим, чем пациенс или посессор пациенса, или когда присутствуют посессоры и агенса, и пациенса, и оба основных аргумента занимают низкую позицию на иерархиях одушевленности,

агентивности и топикальности — в данном случае использование пассивного залога обусловлено не их иерархической позицией, а соображениями удобства восприятия.

**Ключевые слова**: аргументная структура, залог, пассив, посессивность, севернохантыйский язык, уральские языки

**Благодарности**: Исследование проведено в рамках проекта «Глагольное согласование и залог в севернохантыйском дискурсе», поддержанного Фондом Александра фон Гумбольдта. Я благодарен моим коллегам и соавторам Степану Михайлову и Всеволоду Маслюкову за обсуждение моих данных с точки зрения иерархического кодирования и обвиации. Я также благодарю редакторов спецвыпуска и анонимного рецензента, которые помогли улучшить качество этой статьи. Все оставшиеся ошибки являются моими собственными.

**Для цитирования**: Muravyev N. A. The influence of internal possessors on argument coding in Northern Khanty. *Voprosy Jazykoznanija*, 2024, 5: 65–79.

**DOI**: 10.31857/0373-658X.2024.5.65-79

## Introduction

Any discussion of morphosyntactic coding will primarily focus on the properties of the core clausal arguments, namely S(ole) argument of an intransitive clause, A(gent), or transitive subject, and P(atient), or transitive object, following Dixon's [1972] SAO model. One such property that is rarely if at all considered is whether or not a core argument has an internal possessor. In Northern Khanty (< Ob-Ugric < Uralic), the presence of an internal possessor of A and/or P triggers the use of passive voice, as in example (1) below, which does not straightforwardly follow from the identity and status of either of the two core arguments or the event itself.

(1) *śă̈λta* *mewəλ* *nɛpek-əλ* *pox-əl-ən* *šiwalə-s-i*
then chest paper-POSS.3SG son-POSS.3SG-LOC see-PST-PASS[3SG]
'Then his$_i$ (the old man's) son saw his$_i$ chest paper.' (NHC[1], The eagle: 18)

This study explores the idea that such use of the passive voice is due to morphological ambiguity between overt unmarked core arguments (here A 'his son' and P 'his chest paper') and their possessors, as well as referential ambiguity of possessive markers. It is suggested that the passive voice resolves this ambiguity by distinctly marking A with an oblique locative case, thus reducing by one the number of unmarked clausal arguments. Based on the corpus data of Kazym Khanty, I determine the exact conditions under which such ambiguity-resolving passives appear.

## 1. Argument coding in Northern Khanty

Northern Khanty belongs to the Ob-Ugric branch of the Uralic language family. It is spoken in the western part of Northern Siberia, in the Khanty-Mansi and Yamalo-Nenets Districts, by approximately 13,900 speakers, according to the All-Russian Population Census of 2020[2]. Case marking in Northern Khanty follows the neutral/accusative alignment [Nikolaeva 1999]. Full noun phrases as core transitive arguments are always unmarked (2), while personal pronouns distinguish between the unmarked nominative form for the S- and A-arguments and the marked accusative form for the P-argument (3).

[1] The example is taken from the Northern Khanty corpus.

[2] Available online at: https://rosstat.gov.ru/vpn/2020.

(2) *pox an šukat-əs*
boy cup break-PST[3SG]
'A boy broke a cup.' (elicited)

(3) *λuw mănti χătśə-s*
he I.ACC hit-PST[3SG]
'He hit me.' (elicited)

Verbs in Northern Khanty obligatorily agree with the S/A-argument in person and number. Transitive verbs show non-obligatory agreement with P in number, determined primarily by information structure and/or givenness [Nikolaeva 2001; Koshkareva 2002; Dalrymple, Nikolaeva 2011; É. Kiss 2019; Däbritz 2020; Muravyev 2023]. In typological literature, this phenomenon is referred to as differential object indexing, see [Iemmolo 2011; Haig 2018]. Morphologically, verbs that index only the S/A-argument are inflected for subjective conjugation, whereas verbs indexing both A and P arguments take a separate objective conjugation paradigm. Compare a subjective form in a clause with a new and/or focused P-argument (4a) and an objective form with a given and/or topical P-argument (4b). The person-number affix follows the tense marker.

(4) a. *juwan pe:tra re:sk-əs* / **re:sk-əs-li*
John Peter hit-PST[3SG] hit-PST-3SG>SG
[Whom did John hit?] 'John hit PETER.'

b. *luw pe:tra re:sk-əs-li* / **re:sk-əs*
he peter hit-PST-3SG>SG hit-PST[3SG]
[What did John do to Peter?] 'He HIT Peter.' [Nikolaeva 2001: 29–30]

In addition, Northern Khanty has a passive voice that is used whenever the A-argument is not given or topical [Kulonen 1989; Nikolaeva 2001; Koshkareva 2002; É. Kiss 2019]. The P-argument appears in the subject slot and is indexed on the verb by means of the subjective conjugation, while the A-argument is either unexpressed or, if overt, coded as a locative oblique, cf. (5). The passive marker *-a(j)* appears between tense and person-number affixes, with a non-word-final allomorph *-aj* and word-final *-a* [Kaksin 2010: 112–114]. In some Northern dialects, an additional pair of allomorphs, *-i(j)*, is used with vocalic verb stems.

(5) *kalan xoj-na we:l-s-a? — Juwan-na we:l-s-a*
reindeer who-LOC kill-PST-PASS[3SG] john-LOC kill-PST-PASS[3SG]
'Who killed the/a reindeer? — JOHN did.' [Nikolaeva 2001: 25]

Together, the passive voice and both conjugation types form a hierarchical alignment system based on animacy, givenness, and topicality (see [Muravyev 2023]). In languages with hierarchical alignment, the subject slot is accessed by the highest-ranked core argument [Nichols 1992; Thompson 1994; Zúñiga 2006]. Whenever A outranks P, neutral *direct* construction is used, as in example (6a) from Plains Cree with Speech-act participant (SAP) A > 3rd person P. Otherwise *inverse* construction is used, as in (6b) with 3 A > SAP P.

(6) a. *ni-wāpam-ā-w*
1-see-DIR-3
'I see him/her.'

b. *ni-wāpam-ikw-w*
1-see-INV-3
'He/she sees me.' [Zúñiga 2006: 2]

Assuming active voice to correspond to direct and passive to inverse, we get a similar person contrast in Northern Khanty. In SAP A > 3 P contexts, active, as in (7a), is obligatory, because there are no locative forms for personal pronouns, which makes a passive construction impossible. On the other hand, in 3 A > SAP P contexts, passive is strongly preferred (7b), yet active, as in (3), may also occur if A is highly topical.

(7) a. *ma* *waśaj-en* *χătśə-s-əm*
I Vasya-POSS.2SG hit-PST-1SG
'I hit Vasya.'

b. *ma* *waśaj-en-ən* *χătśə-s-ij-əm*
I Vasya-POSS.2SG-LOC hit-PST-PASS-1SG
'Vasya hit me.' (elicited)

There is often a further distinction between "proximate" (discourse-prominent, hierarchically high) and "obviative" (non-prominent) third-person arguments, see Section 2.3 for further details. For example, (8a) and (8b) from Plains Cree illustrate a proximate- and obviative-marked 'dog', respectively.

(8) a. *niwapamaw* *atim*
see(1–3) dog(3)
'I see the dog.'

b. *pakamahwew* *napew* *atimwa*
hit(3–3') man(3) dog(3')
'The man hits the dog.' [Aissen 1997: 707]

In most dialects of Northern Khanty, there are no formal distinctions in the domain of third person. However, Kazym Khanty employs a 2SG possessive affix (cf. (9a)) as a salient article, which specifically marks prominent or "proximate" arguments [Mikhailov (to appear)]. This salient article follows the proximate uniqueness principle proposed in [Goddard 1990] and [Aissen 1997; 2001] for languages with obviation marking. For instance, in (9b), only one noun phrase is compatible with the proximate reading of the 2SG possessive marker.

(9) a. *amp-en* *λow* *šuwəλ-əs*
dog-POSS.2SG bone see-PST[3SG]
'The dog saw the bone.'

b. *amp-en* *λow-en* *šuwəλ-əs*
dog-POSS.2SG bone-POSS.2SG see-PST[3SG]
'The dog saw your/*the bone.' (elicited)

## 2. Passive and ambiguity resolution

Passive constructions in Northern Khanty are highly frequent in texts and have a variety of uses, some of which may seem counterintuitive at first. Consider the following pair of examples:

(10) *mɔś* *pɔxatur* *imij-a* *mă-s* *ar* *sɔrńi* *wox,* *pox-lɛ*
Moshch hero woman-DAT give-PST[3SG] much gold money boy-DIM
*wu-s-λe* *pa* *joxλi* *pɛλa* *măn-ti* *pit-əs*
take-PST-3SG>SG ADD backwards to go-NFIN.NPST AUX-PST[3SG]
'The Moshch hero gave the woman a lot of gold and money, took her son and went back.' (NHC, The hero: 29)

(11) *śi* *kut-ən* *ime-λ* *nox* *wɛrλ-əs* *pa* *jɔš-əλ*
DEM gap-LOC wife-POSS.3SG up wake-PST[3SG] ADD hand-POSS.3SG
*jira* *katəλ-s-a*
to.side grab-PST-PASS[3SG]
'Meanwhile his wife woke up and grabbed his hand (lit. "his hand was grabbed").' (NHC, Three wise words: 70)

In (10), the same actor is expressed throughout the passage as the subject. In contrast, in (11), the actor is first mentioned as the subject but then demoted in a passive construction. Typologically, A-arguments performing successive actions normally continue to be expressed

in the subject slot (see [Givón (ed.) 1983]), as in (10). Thus, the use of the passive in (11) seems out of place. The only obvious difference between the contexts is the presence of an anaphoric third-person possessor 'husband' in the nominal expressions 'his wife' and 'his hand' in the latter example. One may wonder what the link is between expressing a possessor and using the passive voice.

The idea I argue for in this paper is that the passive, as seen in (11), is used to resolve ambiguity, similar to the use of the inverse in Tlapanec to disambiguate between subjects and possessors [Wichmann 2007]. Northern Khanty tends to avoid overt expression of core arguments and instead uses zero-anaphora or pronominal expressions. Combined with the presence of a possessor in (11), this creates ambiguity: it is unclear, especially without contextual support, who grabbed whose hand. The passive efficiently resolves this ambiguity by a) demoting the agent 'wife' so that it can no longer be interpreted as the antecedent of the third-person possessive marker on 'hand', and b) keeping NPs with coindexed third-person possessors in the subject slot to ensure that their possessor is interpreted as the same referent, 'husband'.

If this analysis is correct, we would expect the passive to resolve ambiguity in the presence of a possessor in most or all of the available corpus data. My research questions are as follows:

1. **Does passive resolve ambiguity between A, P, and possessors?**
2. **In what contexts does the ambiguity resolving passive apply?**

The first question can be answered by examining a sample of transitive clauses with possessors and determining the ratio of passive to active examples. If all or nearly all examples are passive, then the passive can be viewed as a dedicated means of ambiguity resolution. Otherwise, the passive is just one among several ambiguity-resolving means, and its use is restricted to certain contexts, which the second question aims to identify. Identifying these contexts requires an in-depth analysis of the event properties and all participants involved.

My general proposal can be formulated as follows:

> **Under conditions that require morphological ambiguity resolution in clauses simultaneously containing a) core arguments A and P, and b) possessor(s) of A and/or P, such ambiguity is fully or partially resolved by passivization, which demotes A to a locative oblique.**

The main research objective was to identify the exact conditions for morphological ambiguity resolution. The following three hypotheses were considered:

- **Processing efficiency hypothesis**: Obligatory A-demotion and passivization occur whenever a clause contains more than two participants from the set {A, P, PsrA, PsrP}.
- **Possessor affectedness hypothesis**: Obligatory A-demotion and passivization occur whenever P's possessor is affected by the event and needs to be morphosyntactically integrated into the proposition.
- **Proximate hypothesis**: Obligatory A-demotion and passivization occur whenever either P or its possessor is more prominent in the current discourse than A.

To test these hypotheses, all 1,035 transitive clauses, including 350 examples with possessors, were extracted from the Kazym dialect subcorpus of the unpublished Northern Khanty corpus (42,174 tokens), compiled by Egor Kashkin. For general purposes and to test the first hypothesis, the corpus examples were manually annotated for active and passive forms of the verb, the presence/absence of possessors on A and P, their formal expression (i.e., full NP or person-number index on the head), and whether they correspond to A, P, a speech-act participant, or another discourse referent. To test the second hypothesis, the examples were annotated according to whether the possessor of P is affected by the event. Annotation for the third hypothesis included parameters that are typologically relevant for hierarchical systems: person

(1, 2, 3), animacy (human, animal, inanimate), definiteness (definite, indefinite specific, indefinite non-specific), and topicality (topical > non-topical). Additionally, some field examples collected in the village of Kazym (Russia) in 2022 were used for illustration purposes.

## 2.1. Processing efficiency hypothesis

According to this hypothesis, the passive voice occurs in transitive clauses when A and P are present along with a possessor of either one or both, without any further conditions. The underlying assumption is that, since neither of the non-pronominal participants is case-marked, the simultaneous presence of these participants creates parsing difficulties. In his discussion of efficiency and complexity in English relative clauses, J. Hawkins [2004] demonstrates that similar parsing difficulties can be addressed by diversifying the coding in a clause, which he calls the "Maximize On-Line Processing" principle.

For example, in a "garden-path sentence" (12a), the hearer cannot unambiguously assign a grammatical function to the NP *the boy* before the sentence is finished. In contrast, (12b) resolves the issue by explicitly introducing the dependent clause with a complementizer *that*.

(12) [Hawkins 2004: 51]
a. *I* $_{VP}$[*realize* $_{S2}$[*the boy knows the answer*]].
b. *I* $_{VP}$[*realize* $_{S2}$[*that the boy knows the answer*]].

In Northern Khanty, the passive voice appears to serve the same purpose by distinctly coding A as a locative oblique, as in (13) below. This allows the hearer to process the sentence as early as possible, avoiding erroneous role assignments and referential conflicts. First, such use of passive signals that neither of the unmarked participants is the agent. Second, it helps to avoid confusion of the antecedents of the two possessive markers: the former, on A *xǫj-əλ-ən* 'man', referring to the topical set of characters, and the latter, on P *tɛpəl-əλ* 'belly', to the overtly expressed possessor *sɔrt* 'pike'. This is similar to the concept of precluding referential conflict [Fedorova 2014] and the idea of passive as a "deconflicter" in A. Kibrik's [2011] terms. In the majority of the annotated examples, the passive voice is indeed chosen as predicted.

(13) *śăλta* *in* *pupi* *sox-əp* *xǫj-əλ-ən* *in* *wǫn* *sɔrt*
then DEM bear coat-PROPR man-POSS.3SG-LOC DEM big pike
*tɛpəl-əλ* *pɛlki* *šomərmə-s-i*
belly-POSS.3SG apart crush-PST-PASS[3SG]
'Then the Man in the Bear Coat crushed the belly of the large pike.' (NHC, Younger daughter of the sun: 80)

On the other hand, there are still plenty of examples in the active voice. For instance, A-arguments that are speech-act participants, as 'we' in (14), are never demoted from the subject slot. This can be explained by the fact that first and second persons are easily distinguished from third persons, so no ambiguity occurs.

(14) *šaj* *jańśi* *ɔməs-s-əm,* *šɔš-ət* *jăŋx-mət-s-ət,* *kuš* *muŋ*
tea drink:NFIN.NPST sit-PST-1SG drake-PL walk-PNCT-PST-3SG though we
*λuw* *ńoxe-λ* *ănt* *λɛ-λ-ew*
they meat-POSS.3SG NEG eat-NPST-1SG>SG
'I sat down to drink tea, the drakes flew in, even though we don't eat their meat.' (NHC, A gunless hunter: 10)

However, even third-person A-arguments can sometimes remain in the subject slot in the presence of a possessor, as shown in (15) and (16) below.

(15) *λɔw-əλ* *šǫŋx-s-əλλe*
horse-POSS.3SG kick-PST-3SG>SG
'He$_i$ kicked his$_i$ horse.' (NHC, God given wealth: 19)

(16) *ike-λ* *λaλ'-a* *leśət-s-a,* *ime-λ* *λɔw*
husband-POSS.3SG war-DAT prepare-PST-PASS[3SG] wife-POSS.3SG horse
*pănən* *tǫ-s-λe*
with.oneself bring-PST-3SG>SG
'They sent her husband to war; he took his wife's horse with him.' (NHC, Golden horse: 42)

Examples like (15) with zero A acting on P 'his horse' do not pose serious problems for interpretation, as A and the possessor of P are coreferent, meaning that there are only two participants involved, not three. In such contexts, the active voice is the only available option in the corpus. Note that in a mirror-image setup with A 'his wife' acting on zero P where P and the possessor of A are coreferent, as in (17), only the passive voice is used. Examples of this kind with non-distinct possessors will be briefly mentioned in Section 2.3 but otherwise remain outside the scope of this study.

(17) *joxi* *λoŋ-əs,* *ime-λ-ən* *ińśəs-λ-a*
home enter-PST[3SG] woman-POSS.3SG-LOC ask-NPST-PASS[3SG]
'He$_i$ came home and his$_i$ wife askes him$_i$.' (NHC, The eagle: 55)

Examples like (16) with zero A acting on P 'his wife's horse', on the other hand, do pose a problem for the first hypothesis. They suggest that certain properties of the arguments either block the use of the passive voice or make it unnecessary. In the following sections, I will discuss other factors potentially influencing the use of passive, namely whether P's possessor is affected by the event (see Section 2.2 on the possessor affectedness hypothesis) and whether A is prominent/topical (see Section 2.3 on the proximate hypothesis). Table 1 below presents corpus counts for the processing efficiency hypothesis. All instances of overtly expressed or anaphorically marked possessors of A and P (including cases of multiple embedded possessors, as, e.g., in (16)) were counted as PsrA and PsrP respectively.

*Table 1*

**Active and passive with different sets of arguments**

| | Active | Passive |
|---|---|---|
| **A, P** | 722 | 200 |
| **A, P, PsrP** | 18 | 43 |
| **A, P, PsrA** | 22 | 19 |
| **A, P, PsrA, PsrP** | 1 | 10 |
| **Total with possessors** | 41 | 72 |
| **Total** | 763 | 272 |

In contrast to sentences without possessors of core arguments, in which the active voice clearly dominates in frequency, the presence of possessors gives the passive voice some advantage[3]. The passive voice appears twice as often as the active voice with A, P, and a possessor of P, and ten times more often than the active voice with A, P, and both possessors. However, with A, P, and a possessor of A, the active voice has a slight edge. Hence, the processing efficiency hypothesis needs further elaboration.

[3] The Fisher exact test statistic value is < 0.00001. The result is significant at $p < 0.05$.

## 2.2. Possessor affectedness hypothesis

Since the simultaneous presence of A, P, and possessor(s) alone cannot account for the use of the passive voice, there must be additional factors at play. One potential factor is the affectedness of the possessor, which triggers passivization, as in example (18) below from Japanese. Such construction is usually referred to as adversative passive [Kortlandt 1992; Shibatani 1994], as typically the promoted participant is negatively affected by the event.

(18) JAPANESE [Kortlandt 1992: 99]
*John wa dareka ni ie o yakareta*
John TOP someone DAT house ACC burn.PASS.PST
'John's house was burnt by someone; John was negatively affected by it.'

In this example, John is the owner of the house and suffers the loss of his property. To emphasize the effect of the burning event on the possessor, the verb appears in the passive voice, and the possessor is topicalized. Northern Khanty exhibits a similar passive construction; however, the affected possessor, as 'Masha' in (19) and the speaker in (20) below, remains internal and is promoted to subject along with its host.

(19) *mašaj-en χɔt χojat-ən λoŋ-s-a*
Masha-POSS.2SG house who.INDEF-LOC enter-PST-PASS[3SG]
'Masha's house was entered by someone.' (elicited)

(20) *xɔt-ɛm xojat-ən nox wu-s-i*
house-POSS.1SG who.INDEF-LOC up take-PST-PASS[3SG]
'My house has been taken by someone.' (NHC, Little chipmunk: 13)

However, examples of the active voice with an affected P's possessor, such as (21) below with zero A acting on P 'the front door of the Vein-making woman's house', are also found in the corpus, which questions the possessor affectedness hypothesis.

(21) *meń kim ɛt-əs, śiw mă̆n-əs pa*
daughter-in-law out appear-PST[3SG] DEM.LAT go-PST[3SG] ADD
*λɔn_wɛrti_ime-n xɔt ɔw-əλ tut-ən*
vein_making_woman-POSS.2SG house door-POSS.3SG fire-LOC
*pon-s-əλλe*
put-PST-3SG>SG
'The daughter-in-law came out, went there and set the front door of the Vein-making woman's house on fire.' (NHC, Vein-making woman: 11)

Moreover, in both active (see (16) repeated below as (22)) and passive (23) examples with possessed Ps 'his wife's horse' and 'the son of those man and woman' respectively, the possessor of P is not affected by the event. Examples with A, P, and A's possessor but no P's possessor are also left unaccounted for.

(22) *ike-λ λaλ'-a leśət-s-a, ime-λ λɔw*
husband-POSS.3SG war-DAT prepare-PST-PASS[3SG] wife-POSS.3SG horse
*pă̆nən tǫ-s-λe*
with.oneself bring-PST-3SG>SG
'They sent her husband to war; he took his wife's horse with him.' (NHC, Golden horse: 42)

(23) *śă̆λta ime-ŋən–ike-ŋən pox śiw śi wɔx-s-a*
then woman-DU–man-DU boy DEM.LAT FOC call-PST-PASS[3SG]
'Then they called the son of those man and woman.' (NHC, Younger daughter of the sun: 46)

Table 2 below presents corpus counts for the possessor affectedness hypothesis.

*Table 2*

**Active and passive with different sets of arguments with an affected and a non-affected possessor of P**

| | PsrP affected | | PsrP not affected | |
|---|---|---|---|---|
| | **Active** | **Passive** | **Active** | **Passive** |
| **A, P, PsrP** | 8 | 23 | 11 | 11 |
| **A, P, PsrA** | — | — | 22 | 19 |
| **A, P, PsrA, PsrP** | 1 | 3 | 0 | 6 |
| **Total** | 9 | 26 | 33 | 36 |

The passive voice does correlate clearly with the affectedness of the P's possessor[4]. It appears three times more often with A, P, and the possessor of P, as well as with A, P, and both possessors. Yet, many examples without an affected possessor of P or without such a possessor are left unexplained by the possessor affectedness hypothesis. Hence, other conditions should also be taken into consideration.

## 2.3. Proximate hypothesis

Another potential factor is the relative ranking of A, P, and possessor(s) in terms of discourse prominence. Here I understand prominence in the sense of [von Heusinger, Schumacher 2019: 119] as "a relational property that singles out one element from a set of elements of equal type and structure". Based on the conventions for describing languages with hierarchical alignment, I will distinguish two prominence statuses: proximate, or discourse-prominent, and obviative, or non-prominent. Proximate is the highest ranked argument among all the arguments in a clause with respect to the hierarchies of animacy, topicality, agentivity and with possessors preferred over possessees [Aissen 2001: 6]. Proximate assignment obeys the Proximate Uniqueness Constraint stating that there can only be one proximate argument per obviation span which is minimally a clause [Ibid.: 12]. All other arguments are assigned obviative status.

In languages with hierarchical alignment, possessors are often treated on par with core arguments. Possessors can be either proximate or obviative, and proximate possessors have distinct coding from obviative ones [Aissen 1997]. In some languages, proximate possessors can even compete with A and P for the subject slot [Nikolaeva et al. 2019]. In a closely related Uralic language, Tundra Nenets, proximate possessors show indexing on the head, whereas obviative ones do not [Bárány, Nikolaeva 2019; Nikolaeva, Bárány 2019]. Compare (24), where the previously mentioned referent 'Wera' is assigned proximate status and is thus obligatorily indexed as a possessor, with (25), where the new referent 'Petya' is assigned obviative status with no indexing on the head.

Tundra Nenets

(24) [Why did Wera hit you?]

| *mәń˚* | *Wera-h* | *mašina-m*$^{*?}$*(-ta)* | *taxabta˚-dәm-ć* |
|---|---|---|---|
| I | Wera-GEN | car-ACC-3SG | break-1SG-PST |

'I broke Wera's car.' [Nikolaeva, Bárány 2019: 232–233]

[4] The Fisher exact test statistic value is 0.0356. The result is significant at $p < 0.05$.

(25) [Where is your cup?]
*Pet'a-h* *weńako(-*?da)* *taxabta°-da-ś°*
Petya-GEN dog-3SG break-3SG>SG-PST
'Petya's dog broke it.' [Ibid.]

Similar contrast is observed in Kazym Khanty, cf. *pet'aj-en xot-əλ* [Petya-POSS.2SG house-POSS.3SG] 'Petya's house' where the indexed possessor 'Petya' is proximate and *pet'aj-en xot* [Petya-POSS.2SG house] where it is non-indexed and obviative. Moreover, proximate possessors do participate in the hierarchical coding system alongside core arguments. Thus, settings with obviative A acting upon obviative P with a proximate possessor in almost all of the examples lead to passivization. In both examples below, passive is used, because the main characters appear as proximate possessors of obviative P-arguments, 'feet' (26) and 'wife' (27), while A-arguments 'those people' and 'Seven Canopies, Big Canopy' are obviative as well.

(26) *kur-λ-aλ* *xɔšməλt-λ-aj-ət,* *isa* *potər*
foot-PL-POSS.3SG warm.up-NPST-PASS-3PL UQ speech
*oš-λ* *ăntǫm*
mind-POSS.3SG NEG.EX
'(Those people) are warming up her feet, she can't speak at all (from fear).' (NHC, Woman with children abandoned: 67)

(27) [When he traveled, he used to leave his heart at home. While he was once driving through the forest, a man named Seven Canopies, Big Canopy, comes to his house.]
*ime-λ* *śi* *λipət-s-a*
wife-POSS.3SG FOC pester-PST-PASS[3SG]
'He clung to his wife [— Where is his heart?].' (NHC, Younger daughter of the sun: 56)

In settings where an obviative A with an obviative possessor acts upon a proximate P, as shown in (28) below, passive also occurs without exceptions. Here P 'the small box' is assigned proximate status because it is more topical in the current stretch of discourse than the possessor of A, 'father'.

(28) *pox-əλ-ən* *aj* *λaraś* *nox* *punš-s-a*
boy-POSS.3SG-LOC small box up open-PST-PASS[3SG]
'His son opened the small box.' (NHC, The eagle: 16)

On the other hand, contexts where a proximate A acts upon P with an obviative possessor, as illustrated in (29) below, consistently demonstrated the use of the active voice in all attested cases. The obviative status of the possessor 'husband' is supported by the fact that it is not indexed on the head noun 'heart'. Therefore, the higher prominence of A compared to the possessor naturally explains the preference for the active voice over the passive.

(29) *in* *nɛ* *pakən-əλ* *pa* *śi* *mă-λ-λe*
DEM woman be.scared-NPST[3SG] ADD FOC give-NPST-3SG>SG
*ike-λ* *săm*
man-POSS.3SG heart
'The woman gets scared and gives her husband's heart.' (NHC, Younger daughter of the sun: 40)

Importantly, examples (15) and (17) with non-distinct possessors discussed earlier in Section 2.1 also fit naturally into this proposal. The obligatory use of the active voice when P is possessed by A and the obligatory use of the passive voice when A is possessed by P are due to the possessor always being more prominent than its possession, as described by the principle "possessor is proximate, possessee is obviative" [Aissen 1997].

One problematic context arises in sentences where both A and P have possessors. Under a hierarchical analysis, both A and P in such cases should be assigned obviative status, hence

no preference for either active or passive voice is expected. However, in reality, we observe a surprising dominance in the use of the passive voice. There are isolated examples in the corpus where there are two distinct possessors, with only one being proximate, as in (30) where the possessor of P 'its sound' is proximate and A 'our adults' has a first-person possessor, which has no obviation status, since speech-act participants do not participate in obviation [Goddard 1990; Aissen 1997; 2001].

(30) *măn-əm* *sij-λ* *wǫn* *ot-λ-am-n* *śi* *xǫλ-s-a,*
go-NFIN.PST sound-POSS.3SG big thing-PL-POSS.1SG-LOC FOC hear-PST-PASS[3SG]
*ma* *isa* *xoλ* *pakən-s-əm*
I UQ away be.scared-PST-1SG
'The noise with which she flew away was heard by adults; I was completely scared.' (NHC, The tackle: 22)

However, in the majority of examples like sentence (1) repeated below with A 'his son' acting on P 'his chest paper', A and P share the same proximate anaphorically expressed possessor, and in this case, passive invariably occurs.

(31) *śăλta* *mewəλ* *nɛpek-əλ* *pox-əλ-ən* *šiwaλə-s-i*
then chest paper-POSS.3SG son-POSS.3SG-LOC see-PST-PASS[3SG]
'Then his$_i$ (the old man's) son saw his$_i$ chest paper.' (NHC, The eagle: 18)

This results in an asymmetrical situation where a possessed P is promoted to subject whenever A is obviative, whether possessed or not, while a possessed A can fill the subject slot only if P is obviative and has no proximate possessor. One possible explanation for this asymmetry, which again involves the processing efficiency hypothesis, is that when a clause contains a full set of A, P, and their possessors, there is no clearly proximate, zero, or pronominal expression of A. It becomes too difficult to keep track of all the roles and possessive relations of the participants; hence A is demoted to oblique via passivization. Table 3 presents corpus counts for the proximate hypothesis. Relative prominence of arguments and possessors was assessed based on their respective positions on the animacy, agentivity, and topicality hierarchies.

*Table 3*

**Active and passive with different sets of arguments and varying prominence of the arguments**

| | A / PsrA > P / PsrP | | A / PsrA < P / PsrP | | PsrA = PsrP | |
|---|---|---|---|---|---|---|
| | **Active** | **Passive** | **Active** | **Passive** | **Active** | **Passive** |
| **A, P, PsrP** | 17 | 0 | 1 | 39 | — | — |
| **A, P, PsrA** | 20 | 0 | 0 | 18 | — | — |
| **A, P, PsrA, PsrP** | 2 | 0 | 0 | 1 | 0 | 9 |
| **Total** | 39 | 0 | 1 | 58 | 0 | 9 |

The proximate hypothesis successfully accounts for almost all of the available data[5]. Active voice consistently appears when A is hierarchically higher than P or its possessor, and the use of passive voice when A is lower than P or its possessor is also almost without exception. Hence, this hypothesis is accepted. The only category of examples not accounted for are sentences with both hierarchically low A and P where only passive voice is used.

[5] The Fisher exact test statistic value is < 0.00001. The result is significant at $p < 0.05$.

## 3. Summary and discussion

The Northern Khanty data presented here demonstrates that, although possessors are not verbal arguments, they still belong to the morphosyntactic core of the system. Similar to A- and P-arguments, they are not marked for case and are indexed for person and number on the head, with the distinction that this indexing occurs inside the noun phrase rather than on the predicate. Functionally, treating possessors in this manner serves the purposes of discourse status tracking by signalling that the current core arguments are linked to a more prominent discourse referent.

On the formal side, however, this treatment of possessors creates ambiguity that is challenging to resolve without proper morphological support. As demonstrated above, this role is fulfilled by the passive voice. First, in the passive construction A is coded as a locative oblique, which clearly distinguishes it from other unmarked noun phrases. Second, even in the absence of an overt A, passive form of the verb signals that none of the unmarked noun phrases is the A-argument. Third, by demoting A from the subject slot, the anaphoric third-person possessor of P can no longer have A as its antecedent.

However, the decision of whether to passivize the clause or not is not solely based on formal considerations when ambiguity arises, but rather depends on the relative prominence of A, P, and their possessors. A prominent A-argument retains the subject position regardless of the presence or prominence of the possessor of the P-argument. If A is not prominent and the possessor of P is prominent, then A is demoted. The prominence of the possessor on A does not seem to play a role, except to signal that A itself is non-prominent and should be demoted when P is prominent.

The treatment of possessors in the system of Northern Khanty raises several questions. First, why is disambiguation achieved through passivization? The answer to this question lies in the central role of topicality in the morphosyntax of Northern Khanty, in contrast to the role-based morphosyntax found in many European languages. In languages where core arguments are distinguished by case marking, such as German or Russian, core arguments and possessors are coded with distinct cases or postpositions depending on their roles. Prominent possessors can also be integrated into the clause as participants in their own right, for example, receiving the dative case (see, e.g., [Shibatani 1994]).

In Northern Khanty, participants are primarily distinguished by whether they are topical or not. If a topical participant is not A, it is then promoted to the subject slot through passivization. Unlike languages with dedicated possessor case markings or postpositions, dative in Northern Khanty appears limited to marking ditransitive recipients and directional participants. Therefore, a possessor can only be indirectly promoted through the promotion of its host, the P-argument, via passivization.

Second, why do we need a broader notion of discourse prominence for Northern Khanty rather than simply adopt the idea of discourse topicality and topic maintenance[6] proposed in [Givón (ed.) 1983] and applied in [Filchenko 2012; Sosa 2017] to Eastern Khanty and by [Koshkareva 2002] to Northern Khanty data? Corpus data show that, although this idea is in most cases applicable, there is evidence that discourse topicality is not always the primary factor in subject selection. Consider in (32) below the context of the earlier cited example (29) in more detail:

(32) [When he traveled, he used to leave *his heart* at home. While he was riding through the forest, a man named Seven Canopies, Big Canopy, came to his house. He clung to **his wife**. “Where is his heart? Show me,” said the man named Seven Canopies, Big Canopy. “If you don’t show it, I’ll kill you right here.”]

| ***in*** | ***nɛ*** | *pakən-əλ* | *pa* | *śi* | *mă-λ-λe* | *ike-λ* | *săm* |
|---|---|---|---|---|---|---|---|
| DEM | woman | be.scared-NPST[3SG] | ADD | FOC | give-NPST-3SG>SG | man-POSS.3SG | heart |

‘The woman gets scared and gives her husband’s heart.’ [The man took *the heart* and threw *it* into the waters of the Ob river.] (NHC, Younger daughter of the sun: 40)

[6] I thank the anonymous reviewer for pointing out this competing explanation.

In this passage, four discourse participants are involved: the husband, his heart, the wife, and a sudden visitor named Seven Canopies, Big Canopy. The heart is globally topical, as it is not only repeatedly mentioned in this passage but is also important for the remaining part of the story. The woman and the visitor are both local discourse topics. The visitor, being the main actor throughout the fragment, is more topical than the woman. Since the woman is the least topical out of the three referents, we would expect the verb 'give' in the target sentence to be passivized and the A-argument demoted to oblique, much like in the example (11) discussed earlier. Yet we see an active clause which seems to be motivated by the fact that not only A 'the woman' is to certain extent topical but also that it outranks P 'heart' in animacy and also holds the subject pivot from the immediately preceding clause, hence being broadly discourse prominent rather than simply topical. Thus, the discourse prominence approach has a slightly better explanatory potential, though it remains yet to be systematically tested on the available data.

Third, what prevents prominent A from being demoted in the presence of possessor(s) and under threat of multiple ambiguity? In the majority of cases, prominent A is referred to with zero-anaphora or takes a pronominal form, i.e. personal pronoun or verbal indexing, whereas less prominent P and its possessor need more expressive material and thus are referred to with full NPs. Such a difference between A and other participants prevents the ambiguity, hence no need for the disambiguating passive.

P's possessor can potentially be mistaken for an overt A, as 'Vein-making woman' in example (21) repeated below. However, were it for the topic shift, 'Vein-making woman' would have probably been stressed prosodically, segmentally and/or with word order.

(33) *meńkim ɛt-əs, śiw măn-əs pa*
daughter-in-law out appear-PST[3SG] DEM.LAT go-PST[3SG] ADD
*λɔn_wɛrti_ime-n xɔt ɔw-əλ tut-ən pon-s-əλλe*
vein_making_woman-POSS.2SG house door-POSS.3SG fire-LOC put-PST-3SG>SG
'The daughter-in-law came out, went there and set the front door of the Vein-making woman's house on fire.' (NHC, Vein-making woman: 11)

No clear cases of active examples with a prominent A expressed as full NP and with a possessed P are found in the corpus. There are still cases like sentence (11) repeated below that opened our discussion in Section 2, where both A 'wife' and P's possessor 'husband' are prominent and zero or pronominally expressed.

(34) *śi kut-ən ime-λ nox wɛrλ-əs pa jɔš-əλ*
DEM gap-LOC wife-POSS.3SG up wake-PST[3SG] ADD hand-POSS.3SG
*jira katəλ-s-a*
to.side grab-PST-PASS[3SG]
'Meanwhile his wife woke up and grabbed his hand (lit. "his hand was grabbed").' (NHC, Three wise words: 70)

In such cases, regardless of A's prominence, it is demoted via passivization. This appears to be the only effective strategy to avoid referential ambiguity between zero A and the possessive marking on P's 'hand'. Therefore, applying J. Hawkins's "Maximize On-Line Processing" principle to the choice between active and passive in Northern Khanty is constrained to instances deviating from the canonical setting where A is zero/pronominal and P's possessor is overt. These situations include cases where A is overt and P's possessor is pronominal, or where both A and P's possessor are equally overt or zero/pronominal, posing a high risk of role misassignment and referential conflict.

There are at least two directions for further research. First, it is necessary to expand the investigation into possessors and possessive marking in Northern Khanty and Uralic languages in general. These categories can be examined from a discourse perspective, focusing on anaphora, reference, and narrative structure. Specifically, it would be valuable to explore when a possessive noun phrase is chosen as a referential expression, how anaphoric links between possessors and their antecedents are established, and how topical chains involving possessors are constructed.

Second, gaining a deeper understanding of ambiguity resolution involving possessors requires a comprehensive study not only of morphological cues such as passive voice and objective conjugation, but also of other factors including word order, prosody, and semantics. For instance, it could be hypothesized that overtly expressed possessors are prosodically integrated with their head nouns, lacking their own intonation contour and featuring shorter pauses compared to other constituents, which facilitates their recognition as possessors.

All of the above topics would also be intriguing to explore within the realms of typology, historical linguistics, and language acquisition.

## Conclusion

The study examines transitive clauses that include core arguments and their possessors. The presence of at least one possessor alongside A and P simultaneously creates ambiguity, which is resolved through various means such as passive voice, which is in focus of this investigation. The passive ambiguity-resolving strategy is employed when A is less prominent than either P or its possessor. Additionally, passive is used to resolve ambiguity when both possessors of A and P are present, and both core arguments are hierarchically low, a condition driven by considerations of processing efficiency rather than relative prominence. Affectedness of P's possessor also influences the choice of passive, although it is considered a secondary factor.

### ABBREVIATIONS

1, 2, 3 — 1st, 2nd, 3rd person
ACC — accusative
AUX — auxilliary
ADD — additive
DAT — dative
DEM — demonstrative
DIM — diminutive
DIR — direct
DU — dual
EX — existential
FOC — focal
GEN — genitive
INDEF — indefinite
INV — inverse
LAT — lative
LOC — locative
NEG — negation
NFIN — non-finite
PASS — passive
PL — plural
PNCT — punctive
POSS — possessive
PROPR — proprietive
PsrA — possessor of A
PsrP — possessor of P
PST — past
SG — singular
TOP — topic
UQ — universal quantifier

### REFERENCES

Aissen 1997 — Aissen J. On the syntax of obviation. *Language*, 1997, 73(4): 705–750.
Aissen 2001 — Aissen J. The obviation hierarchy and morphosyntactic markedness. *Linguistica Atlantica*, 2001, 23: 1–34.
Bárány, Nikolaeva 2019 — Bárány A., Nikolaeva I. Possessors in switch-reference. *Glossa*, 2019, 4(1): 1–36.
Dalrymple, Nikolaeva 2011 — Dalrymple M., Nikolaeva I. *Objects and information structure.* Cambridge: Cambridge Univ. Press, 2011.
Däbritz 2020 — Däbritz Ch. L. *Topik, Fokus und Informationsstatus: Modellierung am Material nordwestsibirischer Sprachen*. Berlin: Walter de Gruyter, 2020.
Dixon 1972 — Dixon R. M. W. *The Dyirbal language of North Queensland*. Cambridge: Cambridge Univ. Press, 1972.

É. Kiss 2019 — É. Kiss K. Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive. *Proc. of the workshop "Clause Typing and the Syntax-to Discourse Relation in Head-Final Languages", Allensbach-Freudental, May 15–17, 2019.* Bayer J., Viesel Y. (eds.). Konstanz: Universität Konstanz, 2019, 163–174.

Fedorova 2014 — Fedorova O. The role of potential referential conflict in the choice of a referring expression. *The Russian Journal of Cognitive Science*, 2014, 1(1–2): 6–21.

Filchenko 2012 — Filchenko A. Continuity of information structuring strategies in Eastern Khanty. *Argument Structure and Grammatical Relations. A crosslinguistic typology*. Suihkonen P., Comrie B., Solovyev V. (eds.). Amsterdam: John Benjamins, 2012, 115–131.

Givón (ed.) 1983 — Givón T. (ed.). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam; Philadelphia: John Benjamins, 1983.

Goddard 1990 — Goddard I. Aspects of the topic structure of Fox narratives: Proximate shifts and the use of overt and inflectional NPs. *International Journal of American Linguistics*, 1990, 56: 317–340.

Haig 2018 — Haig G. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. *Linguistics*, 2018, 56(4): 781–818.

Hawkins 2004 — Hawkins J. A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford Univ. Press, 2004.

Iemmolo 2011 — Iemmolo G. *Towards a typological study of differential object marking and differential object indexation*. PhD thesis. Pavia: Univ. of Pavia, 2011.

Kaksin 2010 — Kaksin A. D. *Kazymskii dialekt hantyiskogo yazyka* [Kazym dialect of Khanty]. Khanty-Mansiysk: Ob-Ugric Institute of Applied Researches and Development, 2010.

Kibrik 2011 — Kibrik A. A. *Reference in discourse*. Oxford: Oxford Univ. Press, 2011.

Kortlandt 1992 — Kortlandt F. On the meaning of the Japanese passive. *Acta Linguistica Hafniensia*, 1992, 24(1): 97–108.

Koshkareva 2002 — Koshkareva N. B. Kommunikativnaya paradigma khantyiskogo predlozheniya (na materiale kazymskogo dialekta) [Communicative paradigm of the Khanty sentence (based on the Kazym dialect data)]. *Yazyki korennykh narodov Sibiri*. Iss. 12. Koshkareva N. B., Oktyabr'skaya I. V. (eds.). Novosibirsk: Institute of Philology of the Siberian Branch of the Russian Academy of Sciences, 2002, 29–44.

Kulonen 1989 — Kulonen U. M. *The passive in Ob-Ugrian*. Helsinki: Suomalais-Ugrilainen Seura, 1989.

Mikhailov (to appear) — Mikhailov S. Diagnosing Northern Khanty unpossessives, or how to tell a synchronically independent marker from its diachronic source. *Linguistica Uralica*. To appear.

Muravyev 2023 — Muravyev N. Passive in Kazym Khanty and the interaction of givenness, topicality and animacy. *Linguistica Uralica*, 2023, 59(1): 49–66.

Nichols 1992 — Nichols J. *Linguistic diversity in space and time*. Chicago: Univ. of Chicago Press, 1992.

Nikolaeva 1999 — Nikolaeva I. *Ostyak*. Munich: LINCOM Europa, 1999.

Nikolaeva 2001 — Nikolaeva I. Secondary topic as a relation in information structure. *Linguistics*, 2001, 39(1): 1–50.

Nikolaeva, Bárány 2019 — Nikolaeva I., Bárány A. Proximate possessors. *Prominent internal possessors*. Bárány A., Bond O., Nikolaeva I. (eds.). Oxford: Oxford Univ. Press, 2019, 228–258.

Nikolaeva et al. 2019 — Nikolaeva I., Bárány A., Bond O. Towards a typology of prominent internal possessors. *Prominent internal possessors*. Bárány A., Bond O., Nikolaeva I. (eds.). Oxford: Oxford Univ. Press, 2019, 1–38.

Shibatani 1994 — Shibatani M. An integrational approach to possessor raising, ethical datives, and adversative passives. *Proc. of the 12th Annual Meeting of the Berkeley Linguistics Society*, 1994: 461–486.

Sosa 2017 — Sosa S. *Functions of morphosyntactic alternations, and information flow in Surgut Khanty discourse*. Doctoral diss. Helsinki: Univ. of Helsinki, 2017.

Thompson 1994 — Thompson C. Passive and inverse constructions. *Voice and inversion*. Givón T. (ed.). Amsterdam; Philadelphia: John Benjamins, 1994, 47–63.

von Heusinger, Schumacher 2019 — von Heusinger K., Schumacher P. B. Discourse prominence: Definition and application. *Journal of Pragmatics*, 2019, 154: 117–127.

Wichmann 2007 — Wichmann S. The reference-tracking system of Tlapanec: Between obviation and switch-reference. *Studies in Language*, 2007, 31(4): 801–827.

Zúñiga 2006 — Zúñiga F. *Deixis and alignment: Inverse systems in indigenous languages of the Americas*. Amsterdam; Philadelphia: John Benjamins, 2006.

*Получено / received 08.07.2024* *Принято / accepted 24.09.2024*