Remarks on areal linguistics in the information structure of the Ob-Ugrian languages

Sachiko SOSA – Susanna VIRTANEN

University of Helsinki – sachiko.toguchi@helsinki.fi
University of Helsinki – susanna.s.virtanen@helsinki.fi

1. Introduction

This study combines together two areas of modern linguistics: information structural analysis and areal linguistics. Our aim is to answer the question on whether there is any areal variation of information structure inside the Ob-Ugrian language branch of the Finno-Ugrian language family. The Ob-Ugrian languages are Khanty and Mansi, spoken in Western Siberia. Both of them have several dialects. As Finno-Ugrian languages, Khanty and Mansi have many common features of Uralic-origin, the lack of linguistic gender, for instance, agglutinativity and stem variation in both nouns and verbs. Common features for Ob-Ugrian languages are observed in the use of the dual category (this also appears in the Sami and Samoyedic branches of the language family), personal-type passive conjugation, a low number of noun cases, and use of the possessive verb ('to have'). There are also a number of features adapted from the surrounding languages, and some of them will be presented in this study.

In this introductory section we will first present the theoretical background of the study in 1.1 and 1.2, some earlier studies in 1.3, and finally the aims of our study in detail in 1.4. Further, language contacts of our target languages are presented in Section 2, our research data in Section 3, and the results of our analyses in Section 4. Conclusions and further question are found in Section 5.

1.1. Areal linguistics and linguistic areas

The targets of areal linguistic research are multi-lingual communities, where features have been adapted from one language to one or several others. Areal linguistic studies often deal with features that appear in neigh-

bouring dialects of individual languages, but are absent from the other dialects of the languages to which these dialects belong, or features appearing inside the community but not outside of it (see e.g. Muysken 2008).

The notion of a linguistic area *(Sprachbund)* has been discussed in literature for decades during the current and the previous centuries. For example, Trubetzkoy (1928) defines the distinction between a genetically related language family and Sprachbund type area by emphasising, that inside a Sprachbund there are often syntactic, morphological and phonological similarities, but not systematic sound correspondences or common basic vocabulary (Trubetzkoy 1928: 18). Campbell (2006) pays attention to the difficulty of defining a linguistic area. She refers to several approaches and attempts to provide a comprehensive definition. Finally, after discussing several criteria such as number of languages, number of language families, number of traits, nature of boundaries, and geography, she concludes that it may be impossible to come to a really satisfying conclusion (Campbell 2006: 18). While there is no meaningful distinction between borrowing and areal linguistics, defining a 'linguistic area' is difficult (Campbell 2006: 1). Campbell emphasises, however, that much more important than defining linguistic areas, is the investigation of the facts of linguistic diffusion.

In this study, our emphasis lies not on the definition of a linguistic area, but on the explicitly visible results of convergence inside one specific linguistic area. We concentrate on the Siberian linguistic area, especially on two languages that are genetically closely related to each other but have also been affected by the surrounding languages. The Ob-Ugrian languages belong to the Finno-Ugric language family and cover a great number of features common to all Finno-Ugrian languages. For thousands of years, however, they have also been influenced by other Siberian and European languages surrounding them. Anderson (2006) has discussed the phenomenon of a Siberian linguistic area and divided the Siberian languages over 40 altogether into 10 genetic units. For example, the Ob-Ugric languages comprise their own unit and the Samoyedic languages their own. He also lists a branch of features commonly shared by all Siberian languages. Further, as a secondary areal feature, he mentions the widespread results of the Russian contacts during the recent centuries.

1.2. Information structure in the Ob-Ugrian languages

According to our recent studies (Virtanen 2015, Sosa 2017), both Khanty and Mansi follow the same principle in information structure: pragmatic functions are realized as syntactic functions. The more topical the argument

is, the higher position it occupies in syntactic hierarchy. There is a clear one-to-one correlation between the pragmatic primary topic and the syntactic subject. The variation between active and passive is due to information structure: in case any other semantic function other than agent is the primary topic of the phrase, the phrase is changed into passive. This correlation is relevant in every Ob-Ugric language form, in both Khanty and Mansi, and in every single dialect (for details, see 4.1).

In some language forms, there is also correlation between the pragmatic secondary topic and a marked syntactic object. For example, in Eastern Mansi, focal arguments are realized as unmarked direct objects (DO's) and obliques, whilst topical DO's are referred to both with verb agreement and case marking. This, however, is not to be applied to all Ob-Ugrian dialects; there is variation inside the Ob-Ugric branch. In contrast to Eastern Mansi, topical objects in Northern Mansi and Northern Khanty are not case marked but only referred to with verb agreement.

In other words, there are basic principles common for all Ob-Ugrian language forms, but, with regard to more detailed features, the dialects differ from each other. The differences do not follow language borders. Our aim in the following sections is to describe these differences and present an areal analysis on them.

1.3. Recent studies: Kulonen 1989

Kulonen (1989) has divided the Ob-Ugric area in three sub-areas according to the appearance of the pronominal passive agent. The Western Area includes Northern Mansi, Western Mansi and Northern Khanty. The Middle Area includes Eastern Mansi, Southern Mansi and Southern Khanty. The third area, the Eastern one, consists of Eastern Khanty alone. Kulonen's areal model is demonstrated in Table 1:

Western group	Middle group	Eastern group
Northern Mansi	Eastern Mansi	Eastern Khanty
Western Mansi	Southern Mansi	
Northern Khanty	Southern Khanty	

Table 1: Ob-Ugric areals according to the appearance
of the pronominal passive agent (Kulonen 1989)

One of the aims of this study was to find out, if Kulonen's results can be applied to our data as well. As can be seen in our results, information structure has the same kind of areal variation as the pronominal passive agent, but the division is not exactly the same. With regard to areal differences in in-

formation structure, the Ob-Ugrian language forms cannot be grouped identically to Kulonen's model. We will present our own model of sub-areas in Section 4.

1.4. Aims of the study

Our starting point in this study is our own recent studies on information structure in the Ob-Ugrian languages and their various dialects. By correlating our recent studies (Virtanen 2015, Sosa 2017) we have observed that there are differences in expressing information structure between different Ob-Ugrian language forms. These differences do not follow the language borders: certain features are found both in some forms of Khanty and some forms of Mansi, but not in the others: in other words, with regard to some features Northern Mansi is different form Eastern Mansi but similar to Northern Khanty etc. Our aim is to find out, whether these differences are dependent on areal factors, i.e. if they are caused by the different language contacts experienced in the different areas.

In this study, we will discuss two information structure-related features in the Ob-Ugrian languages: the marking of a syntactic Subject, and the marking of a syntactic DO. Within the Ob-Ugrian branch as a whole, the case marking or unmarkedness of both these functions is due to information structural factors, i.e. the level of topicality of the argument in question. However, there are minor differences in how information structure affects the morphological marking of these functions. Our aim is to examine, whether the model created by Kulonen in 1989 (see 1.3) can be applied to information structure as well or not. By analysing our own data, we determine whether different areas can also be distinguished for Subject and Object morphology, in the same way as Kulonen (1989) presents for the use of the passive. In this study our emphasis is on mapping the differences between different language forms. The deeper reasons for differences – e.g. details of the contact languages – will be topics of our forthcoming studies.

Subject and DO are syntactic functions. According to our recent studies, however, the basis of the Ob-Ugrian information structure is the correlation between syntactic and pragmatic features: variation between different ways of marking the syntactic functions of Subject and DO is based on information structural factors. This is why we are examining marking of syntactic functions as a part of information structural study. Our hypothesis is that the same language contacts that have affected different areas differently, resulting in differences between the sub-areas of the Ob-Ugrian language branch, have caused differences in the way of marking Subjects and DO's as well. In other

236

words, our aim is to find out, whether there is areal diffusion in Subject marking and DO marking, too.

In the following sections, we will present data on several Mansi and Khanty variants. Different variants have different backgrounds of written use: some of them are regularly used in media and literature, and some have only been written by scholars. Even within the same variant, the confirmed transcription does not exist, but the transcription has been varied and changed many times (Salo 2009). Use of IPA is not appropriate these language descriptions either, since Southern Khanty, Eastern and Western Mansis are now extinct, a phonetic description is unrealistic. For these reasons, different transcriptions are used. In this paper, no mutual transcription has been used for presenting the entire data, rather each individual dialect is represented in a transcription form it was originally recorded in or one used in the most recent studies. This means, that several writing traditions appear beside one another: Northern Mansi is written in Cyrillic, Eastern Mansi in an orthography published by Kulonen in 2007, and so on.

2. Language contacts of Khanty and Mansi

The Ob-Ugrian languages Khanty and Mansi are spoken in the Autonomous District of the Khanty and Mansi in the Russian Federation. This is a multi-lingual area, where several minority languages are spoken in addition to the official state language, Russian. During recent centuries and millennia, variants of Khanty and Mansi have been influenced by many other Siberian languages. Both languages have several dialects, and different dialects have undergone different contacts with surrounding languages. The speaking areas of Khanty and Mansi are not two distinct areas, but the different variants of both languages are spoken in various locations inside the vast area of the district (see Map 1). Sometimes the speaking areas of Khanty and Mansi also overlap with each other. The areas where the Khanty and Mansi live are surrounded by several other minorities – and a Russian speaking population of course. Contact languages include both related and non-related languages. The Khanty and Mansi language contacts have, so far, been mainly investigated from a lexical point of departure. In this section we will briefly discuss the most important language contacts of recent centuries and the present situation.

In the case of both Khanty and Mansi, the Russian influence is much newer than others: It has begun in 1600's in Southern Khanty. The earliest contact in the Ob-Ugrian period was an Iranian language in the 900's, but the

exact knowledge of linguistic contact, e.g. area and timing, is not 100 % clear. After the Iranian contact, the Western and Northern variants of both Khanty and Mansi have been in contact with speakers of Komi languages, at first around the 900's and later in the 1400's. The early Russian loanwords in Western and Northern Mansi and Northern Khanty have also arrived via Komi (Kálmán 1961: 29–36). Even today, some old Komis speak both Komi and Khanty in the region of Ovgort.
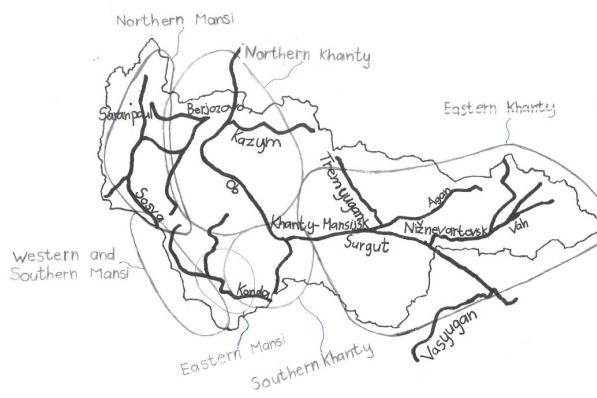
Map 1: Areas where Khanty and Mansi are spoken. (Map by Aura Torri, 2018.)

Sourthern Khanty came into contact with Turkic language forms in 1300's. Most loan words from Turkic are found in Southern Khanty because Eastern and Northern Khanty did not have contact with Turkic. Eastern and Northern Khanty have used loan words from Komi for corresponding concepts. Common Turkic loan words have spread through Southern Khanty and Southern and Eastern Mansi (Abondolo 1989, Toivonen 1945, Steinitz 1961).

Another important contact is Nenets. Steinitz (1959) found that Northern Khanty has 119 loan words from Tundra Nenets. Seven in Northern Mansi, especially the reindeer-breeding terminology has been influenced by the Nenets language. Eastern Khanty has been in contact with Forest Nenets, but their language contact has not been researched well yet (Verbov 1936, Toulouze 2003). Selkup also has been in contact with Khanty, however their contact level is not as intensive as other contacts.

To summarize in brief, each area has its own contact history. In addition to the common younger contact with Russian, the Western variants of both languages are mainly affected by Komi, whilst the Southern ones are effected by Turkic languages, and the Eastern variants of Khanty by Samoyedic languages.

3. Research data

Our data consist of written and audio materials for several Khanty and Mansi dialects. The content of the material is strongly attributed to our recent research projects. As one of the authors has written her PhD on Eastern Mansi, and the other on Eastern Khanty, the corpora used in these studies constitute the basis of our current data: 1200 written entries in Eastern Mansi and around 300 minutes of audio-recorded narrative in Surgut Khanty from Eastern Khanty. In addition, we have gathered smaller corpuses on the remaining dialects: 500 entries on Northern Mansi, 400 entries on Western Mansi and 500 entries on Southern Mansi. Our new Khanty data consist of 100 entries on Southern Khanty and 430 entries on Northern Khanty.

Our Eastern Mansi data have been gathered from the almost only written resource we have – the folkloric collections of Kannisto and Munkácsi. For Northern Mansi we have used the largest up-to-date resource available: *Lūimā Sēripos* is the only Mansi newspaper in the world, published twice a month in Khanty-Mansijsk. Southern Mansi data is from the online-corpus of a Hungarian colleague, Norbert Szilágyi: this data has also been gathered form the Kannisto's and Munkácsi's collections, but it has been transcribed and edited by Szilágyi. Western Mansi data derive from the online corpus on Ob-Babel project: it too has been gathered from Kannisto's and Munkácsi's collections.

Eastern Khanty data mainly consist of the same data used in Sosa's dissertation. They consist of the folkloric collections of, e.g. Csepregi 1998, Csepregi and Sosa 2009, and Sosa's own collection. Southern Khanty data is gathered from the only written resource we have as Paasonen and Vértes (1980). Northern Khanty data is gathered from the online-corpus of two projects on Ob-Ugric languages as *Ob-Ugric languages, Ob-Ugric database* which were provided by Ludwig-Maximilians Universität München. Northern Khanty data base consists of e.g. Rédei (1968), Moldanov (2001) and Solovar (1995).

4. Areal differences in Ob-Ugrian information structure

In this section, we will present the main results of our analysis. In brief, our data show that the basic principles of information structure are identical in the whole Ob-Ugric branch, but there are areal differences on a micro-level: the differences do not follow language borders but the borders of areal sub-groups. The division of sub-groups differ in detail from the division made by Kulonen (1998), but the principle is the same: the diversity of mor-

phological devices used in Subject and Object marking varies inside the Ob-Ugrian branch, providing three different areal groups.

4.1. Common information structural features for all Ob-Ugrian languages forms

There are two very characteristic features common for all Ob-Ugrian language forms:

1) Both Khanty and Mansi are configurational languages: information structural functions are primarily expressed by variation between different syntactic structures, e.g. between active and passive.

2) In addition, the variation between two verb conjugation paradigms is connected to information structural features: the so called object conjugation is used for marking topicality of the DO in the whole Ob-Ugric branch (see Skribnik 2001, Virtanen 2015, Sosa 2017).

In all Khanty and Mansi dialects the variation between active and passive, and the variation between different three-participant constructions, are based on information structure: there is a correlation between pragmatic and syntactic functions. The most topical element always occupies the syntactic role of Subject. This causes variation between active and passive. As demonstrated in (1) concerning Northern Mansi, for example, a sentence with the semantic Agent as the primary topic is realized in the active. In this case, *people* is the topical, already known element in the text, and the sentence provides new information on their activities:

(1) **Мāхум** *тēлы* *ялп-ыӈ* *хōтал-ыт* *сыс*
  people winter saint-ADJ day-PL POSTP

  тув *ёхт-ал-ас-ыт.*
  there come-DER-PST-PL
  '**People** came there to spend their winter holiday.' (LS 1/2018)

As demonstrated in (2), a sentence with any other semantic function as the primary topic, is realized in the passive. In this example, the text has told about the tourists, and in this particular sentence, new information is provided about someone else taking them for a sleigh ride:

(2) *Тувыл* **мӯй** **хōтпа-т** *сāл-ы* *сун-ыл*
  then guest person-PL reindeer-ADJ sleigh-INSTR

  ос *āмп-ыӈ* *суныл* *тот-ыгл-авē-с-ыт.*
  and dog-ADJ sleigh-INSTR take-DER-PASS-PST-PL3
  'Then **the guests** were taken for a ride with a reindeer-sleigh or a dog-sleigh.' (LS 1/2018)

The same kind of variation is found also in Khanty. In all of the Khanty variations, Subject is the primary grammatical role for topic, both in active and passive. The semantic agent is realized as a Subject in an active phrase, when it is the primary topic (example 3 and 4). The following examples are from a Northern Khanty folklore tale. In example 3, the subject *I*, the boy, is the main character of the tale:

(3) **ma** *ki* *noχpit-ʌ-əm,* *χanti* *χǫ,* *χanti* *nɛ,*
   1SG if prevail-PRS-1SG Khanty man, Khanty woman,

   wǫʌəpse-ʌ *ʌaškam-a* *pit-əʌ.*
   life-SG<3SG free-DLAT become-PT.3SG
   'If **I** prevail, the life of the Khanty man and the Khanty woman will be free.'
   (Rédei 1968: 84)

In the same tale, the main character keeps the status of the primary topic in discourse. However, another referent comes into as the semantic agent. The primary topic is still the main character as patient, even though it is a general tendency that the semantic agent has often topical status. In such a case, the primary topic is realized as Subject of passive in order to keep the status of the primary topic. The semantic agent is realized as the agent in a locative in passive structure and functions as focus:

(4) *jalań* *ikij-ən* *šiw* *ńǫrtemə-s-**i.***
   Old.man.Jalan man-LOC there press-PST-PASS**[3SG]**
   '**He** was pressed by the Old Jalan.' (Rédei 1968: 86)

Furthermore, variation between different three-participant constructions is also motivated by the pragmatics. In an active three-participant clause, the second most topical element (secondary topic) always occupies the syntactic role of DO. The following examples on Eastern and Northern Khanty demonstrate the division of different ditransitive alignments:

(5a–b) Northern Khanty (Dalrymple – Nikolaeva 2011: 148)

Ma *an* *petra* *elti* *ma-s-em/ma-s-əm.*
1SG cup peter to give-PST.SG<1SG / give-PST.1SG
'I gave a/the cup to Peter.'

Ma *petra* *an-na* *ma-s-em/*ma-s-əm.*
1SG peter cup-LOC give-PST.SG<1SG / give-PST.1SG
'I gave a/the cup to Peter.'

241

(6a–b) Eastern Khanty (Sosa 2017: 118)

ʌüw	mant	kat	quʌə-χən-at	məj
3SG	1SG.ACC	two	fish-DU-INSFIN	give-PST.3SG

'She gave me two fish.'

ʌüw	mantem	kat	quʌə-χən	məj
3SG	1SG.DAT	two	fish-DU	give-PST.3SG

'She gave two fish to me.'

While the Northern Khanty ditransitive structure triggers object conjugation based on the information structure in example 5, the ditransitive structure in Eastern Khanty can trigger both subjective and objective conjugation.

The same system can be found in any other Ob-Ugrian dialect. In other words, the main principles are the same in every Ob-Ugrian language variant. However, some details, such as the marking of topical and focal DO's, attest to variation within the Ob-Ugrian area. The differences are discussed in 4.2 and 4.3.

4.2. Areal differences in Subject marking

Regarding Subject marking, the Ob-Ugrian area can be divided into two groups: one of them represents canonical Subject marking, and the other one non-canonical Subject marking. Eastern Khanty alone belongs to the second group: all the other language forms carry the feature of canonical Subject marking.

4.2.1 Canonical Subject marking

Northern and Southern Khanty as well as all variations of Mansi have a canonical nominative subject: both nominal and pronominal Subjects are unmarked or only referred to with a verb ending. This concerns both active and passive sentences: in both of them, the predicate verb correlates with the syntactic Subject. Examples of every language variant are listed below. Often the syntactic subject is not explicitly marked at all: it is only encoded to the predicate verb (see the passive Eastern Mansi and Western Mansi below).

Northern Khanty (Kazym)

Active	śi	jŏχan-ət	a:r	χŭʌ	taj-s-ət.
	That	river-PL	many	fish	have-PST-PL3

'These rivers had many fishes.' (Rédei 1968: 32)

Passive *aʌ*   *šiwaʌə-s-ij-ət*   *pa*   *mɔjpər-ət-ən.*
                NEG   notice-PST-PASS-3PL   other   bear-PL-LOC
                'So that they would not be noticed by the bear.' (Rédei 1968: 40)

Eastern Khanty (Surgut)

Active *Ma*   *nüŋ-at*   *nik*   *ťi*   *tärt-ʌ-əm.*
              1SG   SG-ACC   down   PTCL   grill-PRS-1SG
              'I will roast you.' (Csepregi 1998: 68)

Passive *Qŏʌtåɣiʌ*   *ʌüw*   *järnas-at*   *wär-ʌ-ojəm.*
               Tomorrow   PTCL   cloth-INSFIN   make-PRS-PASS.1SG
               'The clothes will be made for me tomorrow.' (Csepregi 1998: 84)

Southern Khanty (Konda)

Active *Män, ıəɣpax, üš xowən wax-en*   *täš-en*
              1SG brother old long   steel-SG<2SG.POSS   ware/goods-SG<2SG.POSS
              rot-a   *att-em*   *pan-em ---*
              ship-LAT   set-PST.SG.1SG   put-PST.SG<1SG
              'A long time ago, my brother and I set and put your money and goods
               to the ship.' (Paasonen – Vértes 1980: 24)

Passive *Ux-əŋ*   *ıəpe sem-əŋ*   *ıəpə-nə*   *ıəpə-sox-ta*   *kew-əŋ*   *pos-ɣən*
              Head-ADJ owl eye-ADJ owl-LOC owl-tear-INF stone-ADJ mitten-DU
              kärə-past-ikəı-ä   *taxəma-ı-ɣən.*
              FL-FL-man-LAT   throw-PASS-PRT.3SG
              'Mittens were thrown to the *kärə-past-iki*-man by the owl.'
              (Paasonen – Vértes 1980: 6)

Eastern Mansi (Konda)

Active *Am*   *nää-n*   *tat-øs-løm*   *nee-l.*
              1SG   2SG-ACC   bring-PST-SG<1SG   woman-INSTR
              'I brought you a wife.' (WV II: 29B)

Passive *Kom-øjäg-nø*   *lyðnk-øl*   *mäj-w-øs.*
               man-DU-LAT   advice-INSTR   give-PASS-PST
              '[He] was given advice by the two men.' (WV III: 9B)

Northern Mansi (Sosva)

Active *Няврам-ыт*   *щāгтым*   *тот*   *хāйт-ыгтӭг-ыт.*
              child-PL   glad   there   run-DER-3PL
              'Children are running cheerfully.' (LS 1/2018)

Passive *Мāхум пустāғыл ēмтн-ǯныл мус тот ōньщ-ав-ет.*
　　　　 people healthy become-DER- until there have-PASS-3PL
　　　　　　　　　　　　 SG<1PL.POSS
'People are kept there until they are healthy again.' (LS 1/2018)

Western Mansi (Vagilsk)

Active *anʲsʲəx weri utt-əs*
　　　 man weir build-PST[3SG]
'The man built a weir.' (WV III: 15)

Passive *anʲsʲəx-nə a:t ta:rtalt-w-əs*
　　　　 man-LAT NEG let.go-PASS-PST
'She was not let to go by the man.' (WV III: 15[1])

Southern Mansi (Tavda)

Active *äm nüǘn ńēl-én meü-lém*
　　　 1SG 2SG.LAT arrow-sg<2sg.poss give-SG<1SG
'I will give you the arrow.' (WV III: 16)

Passive *čalkan čil tumlant-əu*
　　　　 carrot continuously steel-PASS
'The carrots are stolen all the time.' (WV III: 16)

4.2.2 Non-canonical Subject marking in Eastern Khanty

In addition to the canonical subject marking mentioned above, in Eastern Khanty discourse, even a locative subject appears. This locative subject structure is quite rare in Surgut Khanty, however in more Eastern variations, such as Vah and Vasjugan Khanty, it is quite common. In fact, in some statistics, it is more common than the passive structure which also has a locative agent (Kulonen 1989: 301, Filchenko 2006).

Surgut Khanty:

(7) *ma-nə tŏwə äsλ-em*
　　 1SG-LOC to.there leave-PST.SG<1SG
'I left (it) (there).' (Csepregi 1998: 56)

---

[1] The last number in the reference refers to the number of the text, not a page number. For example: (WV III: 16) – Wogulische Volksdichtung, 3rd Volume, 16th text.

Vasyugan Khanty:

(8) *ɔ̆llə*   *sart*   *rätš*   *män-nə*   *öɣöli-s-im*
    Big   pike   old.man   1SG-LOC   prepare-PST1-SG<1SG
    'I got the big pike ready.' (Filchenko 2010: 399)

The locative subject has been called as "ergative structure" (e.g. Honti 1984, Kulonen 1989, 1991, Ruttkay-Miklián 2002), even though there is no ergative-absolute division in Khanty. In fact, the locative Subject is found also in the intransitive clause in Surgut Khanty discourse. We call this a locative subject structure here (Sosa 2017: 40–43, 182–207). From the perspective of typology, subjects should not be identified based on the traditional morphosyntactic criteria for subject alone (Croft 2003: 14–15, Haspelmath 2010). In the pragmatic analysis of the use of locative subject, it is visible that they are motivated by certain pragmatic conditions in the preceding discourse (Filtchenko 2006, Sosa 2008 and 2017: 182–207).

Eastern Khanty discourse provides the evidence that the locative subject structure has two functions: 1. Returning topic. 2. Emphasised subjectness especially in the relationship between competing topical referent (object) and in the topic shift in discourse. (Sosa 2017: 182–207.)

This exceptional Subject-marking is similar to Tundra Nenets: Tundra Nenets also has locative-inflected Subjects occurring in connection with particular verbs, which demand an experiencer-type Subject (see e.g. Nikolaeva 2014). These Nenets features may relate to the appearance of locative-inflected Subject in Eastern Khanty, which has been in contact with both Nenets forms for a long time. However, this needs more evidence to explain this phenomenon.

4.3. Areal differences in Object Marking

Regarding DO marking, we can distinguish three different zones inside the Ob-Ugrian area:

1) Canonical DO marking: Nominal DO's are never case marked, but the topical ones get verb agreement. Pronominal DO's are case marked.

2) Non-canonical DO marking with two case variants: Topical DO's are case marked and verb marked, whilst focal ones are neither case marked nor verb marked. Pronominal DO's are marked in the same ways as the topical nominal DO's.

3) Non-canonical DO marking with three case variants (Eastern Khanty alone): Topical DO's are verb marked but not case marked, whilst focal ones

are case marked but not verb marked. Pronominal DO's are marked differently from the topical nominal DO's

The division is done regarding the number of noun cases used with nominal and pronominal DO's. Examples of possessive marked nominal DO's will be presented in this study as well, and they cover some exceptional feature: in some language forms, possessive suffixes predominate over case marking. The zones are not identical with the division of sub-areas by Kulonen (1989). We will present the language forms with canonical object marking in 4.3.1, language forms with non-canonical object marking (except Eastern Khanty) is 4.3.2, and non-canonical object marking in Eastern Khanty in 4.3.3.

4.3.1 Canonical Direct Object marking

Canonical DO marking is found in Northern Mansi, Northern Khanty (with the exception of Kazym Khanty, see 4.3.2) and Southern Khanty. In these languages, all nominal DO's are morphologically unmarked. In Northern Mansi and all variations of Khanty, only pronominal objects have accusative marking. Canonical marking, however, concerns only case marking: in all of these language variants, two verb conjugation categories are involved: focal DO's are accompanied by subject conjugation, and topical objects by object conjugation. In other words, topicality of a DO is encoded in the verb, with the exception of personal pronominal object in Eastern Khanty.

Example (9) is from **Northern Mansi:** a focal object is unmarked and accompanied by subjective conjugation:

(9) *Ань балок-т   ōл-нэ   щēмья-т   кол-ыт   те   ёвт-ӗгыт.*
    now container-PL live-PCTP family-PL house-PL PARTIC buy-3PL
    'Now [they] are buying houses for those people living in container houses.'
    (LS 1/2018: 5)

In (10) a topical DO is unmarked but accompanied by the object verb conjugation:

(10) *Тувыл   та   юи=пāлт ущ   олн-ыт   лāкква=уртса-нӯв.*
     then   that after   just money-PL PREF-share-PL<1PL
     'Just after that we will share the money.' (LS 1/2018: 3)

In (11) the topical DO 'speech' accompanied by an infinitive verb form, an auxiliary, and the nominal object constituent is morphologically unmarked:

(11) *Диктант   хас-нэ   пора-т   потыр   нила   щёс*
     dictation write-PT.PTCL time-LOC   speech four   time

ловиньта-ӈкве   рōв-и.
read-INF   is.allowed-3SG
'It is allowed to read the text four times during the dictation assessment.'
(LS 5/2017)

Regarding nominal objects, Northern and Southern Khanties have only one case, nominative. Only personal pronouns are inflected in the accusative for objects in all variations of Khanty.

In **Khanty**, too, topical objects are accompanied by the object conjugation and focal ones as well. The system of DO, however, is not as simple as in Mansi. At first, we will show the simple canonical object marking. The focal DO triggers the subject conjugation, whereas the topical DO triggers the object conjugation. Examples 12 and 13 are from **Northern Khanty**. In example 12, the object is a focal referent whereas in example 13 the subject 'he' is the main character and the object 'us' is secondarily topical:

(12) *Kim   ɔwxari-ja*   **kawərt-əm   n´oxi pul-ije**   *pon-s-əm.*
Outside courtyard-DLAT cook-PTCP-PRS meat small.piece-DIM put-PST-1SG
'I put **a piece of boiled meat** outside in the courtyard.'
(Solovar 1995:89)

(13) *tɔwi-jən*   *amtəptə-ʌ-ʌe*
spring-LOC   gladden-PT-PL<3SG
'In spring, he gladdens us.' (Nemisova – Kajukova 2007)

In addition to the above mentioned simple canonical marking system, in one variation of the **Northern Khanty**, **Kazym Khanty,** personal pronouns are inflected in the accusative and dative cases. The Kazym Khanty inflectional paradigm of personal pronouns shows two inflectional forms, the so called simple and complex ones.[2] According to Klumpp (2012), the inflectional difference between the two accusatives corresponds to the difference in functions: the complex form functions as a focal object, whereas the simple form as topical (14a-b):

(14a) Context: The protagonist is in a boat on a river, suddenly there appears an armoured hero, steps down to the river and addresses the protagonist:

---

[2] Note, that Kazym Khanty consists of several subvariations: Some subvariations have two inflectional forms in the accusative, some do not.

wŭṇšəpt-e   *man**t***   **pa**   *peḷək-a*   *såra*
take.over-IMP.SG<2SG   1SG.ACC1   other   side-LAT   quickly
'[take[me]TOP quickly over to the other side]FOC!' (OA III: 478–479,
in Klumpp 2012: 365)

(14b) Context: After having taken him over a second hero wants the same. Again, the
main protagonist does as requested. This second hero, finally, jumps ashore and says:

Manti   *λawλ-e*   *tăta!*
1SG.ACC2   wait-IMP.SG<2SG   here
'Wait here for [me]FOC!'

In summary, in Group 1 only pronominal DO's are case marked. Pro-
nominal DO's and topical nominal DO's trigger verb agreement. The results
presented in 4.3.2 are gathered together in Table 2:

	Topical nominal object	**Focal nominal object**	**Pronominal object**
Northern Khanty	Nom. + Obj.c.	Nom. + Subj.c.	Acc. + Obj.c.
Southern Khanty	Nom. + Obj.c.	Nom. + Subj.c.	Acc + Subj.c
Northern Mansi	Nom. + Obj.c.	Nom. + Subj.c.	Acc. + Subj.c.

Table 2: DO coding in Group 1.

4.3.2 Non-canonical Direct Object marking with two case variants

Regarding nominal DO's, non-canonical object marking with two case
variants appears in Eastern Mansi, Southern Mansi and Western Mansi.
These language forms are so called DOM-languages[3] DO's are either un-
marked or case marked, depending on their level of topicality. In addition to
case marking, verb agreement is also present. In many earlier studies, the
referential feature of variation is mentioned to be definiteness. According to
more recent studies, including our own studies, this question is not one of de-
finiteness but topicality (Virtanen 2014: 69–71).

In **Eastern Mansi**, topical DO's are accusative marked and encoded in
the verb, whilst the focal ones are unmarked (nominative) and accompanied
by unmarked verb conjugation (Virtanen 2014, 2015). Example (15) repre-
sents a focal, unmarked object in Eastern Mansi. The object is in the nomi-
native, and not encoded in the verb:

---

[3] DOM = Differential Object Marking, see e.g. Bossong 1986, Aissen 2003.

(15) *jänii*   *lyüüly*   *wöär-s-øm*
      big   mistake   make-PST-1SG
      'I made a big mistake.' (WV I: 39)

In (16) we can see a topical DO in Eastern Mansi. The DO is accusative-marked and encoded in the verb:

(16) *luj=ootr-äg-mø*   *wot-öän*
      down=prince-DU-ACC   call-IMP.DU<2SG
      'Call the princes of the Underworld here!' (WV I: 15)

In (17) we can see a pronominal DO in Eastern Mansi, also inflected in accusative:

(17) *om   nää-n   seeməl mõõ   kårøng täw-nø   jål=täärøt-øs-løm*
      1SG 2SG-ACC black Earth dry.land surface-LAT down=let-PST-SG<1SG
      'I let you to the surface of Earth.' (WV IV: 6)

In **Southern Mansi** as well**,** DO's are either unmarked or accusative-marked. In contrast to Eastern Mansi, there is no direct correlation between topicality and case marking. Honti (1969) mentions three different situations: unmarked indefinite objects, case marked definite objects and unmarked definite objects. He emphasises that the case marking of a DO is not in a one-to-one correlations with its definiteness.

Honti regards the variation as a question of definiteness. As already mentioned, however, later studies have shown that topicality is a more accurate way to express the phenomenon. We still accept Honti's division of three types, but instead of definiteness, we will speak of topicality. As applied to our data, we can talk about an unmarked focal object, case-marked topical objects and unmarked topical objects.

In (18) the DO is focal and for that reason unmarked and accompanied by the Subject conjugation.

(18) *ńokor*   *näü*   *tīwa*   *wās-ən*
      what.kind.of   2SG   miracle   see-PST-2SG
      'What kind of miracle did you see?' (WV III: 19)

In (19) the DO is topical, marked with accusative case and verb agreement.

(19) *konkā̄-me*   *pā̄tī-s-tīl*
      reindeer-ACC   shoot-PST-SG<3SG
      'He shot the reindeer.' (WV III: 19)

In (20), however, the topical object is unmarked and only accompanied by the object conjugation:

(20) *kwark-ət*   *šekwelā̃k-ət*   *äm*   *jil=på̊så̊nt-ēləm*
      crow-PL   magpie   1SG   PREF=drive.up-PL<1SG
      'I drive away the crows and the magpies.' (WV III: 19)

As described above, nominal DO's have two case variants. Pronominal objects, however, are always marked with accusative as in (21):

(21) *ämā̃nmi*   *jíl=pō-s-tān*   *kül*   *küer*   *jü=tōla-s-tān*
      1SG.ACC   PREF=capture-   house   into   PREF =take- PST-SG<3PL
                   PST-SG<3PL
      'They caught me and took into the house.' (WV III: 19)

So far, we have presented language forms, where a nominal DO is either unmarked or marked with the accusative case, which is typologically the prototype case of DO. However, among the Ob-Ugrian language forms, there are also examples of DO's marked by oblique cases. In **Western Mansi**, DO's are either lative-marked or morphologically unmarked. The lative case as an object marker is accompanied by the object conjugation. The following examples (22) and (23) are from Northern Vagilsk.

A focal DO is unmarked and accompanied by subjective conjugation – in the same ways as in Eastern and Southern Mansi.

(22) *torəm*   *nʲowlʲ*   *mi-s*
      god   meat   give-PST
      'The God has given meat.' (WV III: 15)

In (23) a topical DO is marked with lative and accompanied by the object conjugation. In contrast to the Eastern and Southern dialects, a possessive marked noun also includes a case ending.

(23) *kurt* *isʲyø-e:t-nə*   *ke:t-s-tə*   *weri*   *tuʃl-əx*
      third daughter-SG<3SG.POSS-LAT   send-PST-SG<3SG   weir   check-INF
      'He sent his third daughter to check on the weir.' (WV III: 15)

Pronominal DO's, too, are marked with lative as in (24) and (25):

(24) *tæw*   *tanən-nə*   *nuŋk*   *lesʲit-es-æ:n*
      3SG   3PL-LAT   PRVB   heal-PST-SG<3SG
      'She healed them.' (WV III: 15)

(25) *utt-əs-tə*   *tæ:w-əte:n*   *kʷælkən*   *ʃim-nə*
place-PST-SG<3SG   3SG-LAT   floor   middle-LAT
'[She] placed her in the middle of the floor.' (WV III: 15)

Devices for marking topical and focal nominal DO's and pronominal DO's in Group 2 (non-canonical Object Marking with two case variants) are presented in Table 3:

	Topical nominal object	**Focal nominal object**	**Pronominal object**
Eastern Mansi	Acc + Obj.c.	Nom. + Subj.c.	Acc + Obj.c.
Southern Mansi	Acc + Obj.c. / Nom. + Obj.c.	Nom. + Subj.c.	Acc + Obj.c.
Western Mansi	Lat + Obj.c.	Nom. + Subj.c.	Lat. + Obj.c.

Table 3:
Devices of marking topical and focal nominal and pronominal DO's in Group 2.
(Obj.c. = Object Conjugation, Subj.c. = Subject Conjugation)

4.3.3. Non-canonical object marking with three case variants (Eastern Khanty)

Object marking in Eastern Khanty has three case variants: the nominative (noun phrase), accusative (personal pronoun) and instructive-final (as "oblique object"). Here the nominative and accusative objects marking adheres primarily to the same strategy as the canonical object marking mentioned above. The topical object triggers the object conjugation whereas the focal object triggers the subject conjugation. In example 26, the object is accompanied by the subjective conjugation in correlation with its focal status:

(26) *sar*   *ma*   *ker*   *nŏq*   *üʌəm*
Immediately   1SG   oven   up   light.PRT.1SG
'Immediately, I heat an oven.' (Csepregi 1998: 66)

In the following example, the object is accompanied by the object conjugation because of its topicality. In example 27, the subject *Mäŋk iki*-troll tries to eat the little bird. In this tale, the little bird is the main character, but the *Mäŋk iki*-troll is more topical in the local discourse. The topical referent *little bird* triggers the objective conjugation:

(27) *piťəŋkəli*   *ťə*   *kem*   *ńərimtə-təɣ*
little.bird   well   out   bring-PST.SG<3SG
'He (*Mäŋk iki-*) took the little bird out'. (Csepregi 1989: 66)

251

In **Eastern Khanty**, the marking of topicality is not as simple as in other Khanty variants. In contrast to all Mansi variants, for example, the Eastern Khanty object conjugation does not appear with pronominal objects, which are generally topical referent (Sosa 2017: 158-163). In example (28), the pronominal referent *me* is a topical object but triggers the subject conjugation:

(28) ***mant***   *panpə*   *pɪrip-əγ*
     1SG.ACC   and   ask-PST.3SG.
     'And (my father) asked **me**.' (Csepregi 1998: 62)

This feature seems to appear only in Eastern Khanty. In Northern Khanty, personal pronouns can trigger both subject and object conjugation, and in all Mansi variants they are always accompanied by the objective conjugation. Example (29) shows that personal pronouns might or might not trigger the objective conjugation. This example also illustrates that the criteria for choosing the conjugation do not depend on definiteness:

Alignment pair in conjugation (Nikolaeva 1999: 65):

(29) *ma* *naŋ-en*   *wan-s-əm*
     1SG 2SG-ACC   see PST-SG<1SG
     'I saw you.'

(30) *ma*   *naŋ-en*   *wan-s-em*
     1SG   2SG -ACC see-PST-SG<1SG
     'I saw you.'

Other examples from Northern Khanty:

(31) *xunsi*   *naŋ*   *muŋi-luw*   *xals´a*   *wantlze-l-an?*[4]
     When 2SG   1PL-ACC   where see-PT-PL<1PL
     'When will you see us, and where?'
     (Nikolaeva 1999: 60; Pápay 1906, 1908)

---

[4] Nikolaeva (1999: 26) defined temporal category in Northern Khanty as the Non-Past, the Past and the future. However, we distinguish temporal category of Northern Khanty here more simplily as the Past and the Presens, since Khanty has only two morphological markings for tence as -L- as the Presens and future, -s- as the Past. She glosses the examples also using these three catogories. Original example XX is also glossed as Non-Past whereas here we glossed the same sentence as Past.

Objective conjugation + personal pronoun object:

(32) *muŋ   naŋ-en   mojlə-ptə-s-luw*
     We   you-ACC   visit-CAUS-PST-PL<1PL
     'We received you (as guests).' (Nikolaeva 1999: 40)

This difference between Northern and Eastern Khanty also influences the choice of conjugation in ditransitive structure. In Northern Khanty, the object of the ditransitive structure triggers the object conjugation because of its topicality (Example 5b). In Eastern Khanty, however, the object of the ditransitive structure also triggers the subject conjugation regardless of its topicality (Example 33).

Subjective conjugation + personal pronoun object:

(33) *ma   naŋ-en   nem-na   pon-l-əm*
     1SG   2SG-ACC   name-LOC   put-PRS-1SG
     'I will give you a name.' (Nikolaeva 1999: 62; Pápay 1906, 1908)

In the Eastern group, Eastern Khanty, the 'oblique object' appears in discourse, in addition to the aforementioned object markings. This oblique object in the instructive-final case forms the alignment with the alternation of a nominative/accusative object. The oblique objects seem to appear with certain verbs in Eastern Khanty discourse, e.g. *λăŋq-ta* 'to want', *λăγλəqsə-ta* 'to wait', *λejλəγəλ-ta* 'to see well'. Among these verbs, *λăŋq-ta* 'to want' trigger always instructive-final case in the data (Sosa 2017:123).

λăŋq-ta 'to want' triggers instructive-final case. (An interview with an informant, Sosa 2017: 124.)

(34) *aŋki   nüŋ,   **jiŋk-at**   ʌăŋw-ən?***
     Mother   2SG   **water-INSFIN**   want-PST-2SG
     'Mother, did you want [any] water?'

As in the description of the non-canonical Subject, DO's should not be identified on the basis of a the morphosyntactic definition of transitivity. The oblique object can be interpreted as an object based on the semantic principal: with the oblique object structure, the Subject and the oblique object related to agentivity (agent) and affectedness (patient) and perfectivity (in real time). (Hopper and Thompson 1980.) In other words, transitivity is based on semantics and pragmatics also, not only in the indexing of case marking (Iemmolo 2011). Based on the indexing of case marking, the instructive-final is an oblique argument, whereas the nominative/accusative case, which is canonical object marking case, belongs to the category of core arguments. In

253

general, a core argument represents given information and is trackable in discourse, and can also become a discourse topic, whereas an oblique represents new information and is nontrackable (e.g. Laury 1997).

In Eastern Khanty, oblique objects tend to refer to indefinite and/or generic referents which less trackable than canonical objects (Sosa 2017: 121–134). Examples 35a-b are extracted from a fairytale:

Nominative object:

(35a) *T'u*   *jaγ*   ***påri***   *wär-ət.*
      This   people   **feast**   prepare.PST-3PL
      'The people prepare a feast.'

Oblique object:

(35b) *Qŏʌtaγəʌ*   *sär*   *jəm*   ***părij-at***   *wär-ʌ-əttən.*
      Tomorrow   forward   good   **feast-INSFIN**   prepare-PT-2DU
      'The two of you are preparing a large feast tomorrow indeed.'
      (Csepregi 2011: 19; Sosa 2017: 129)

In the above example (34), the referent *påri* 'feast' appears both in the nominative (35a) and in the instructive-final (35b) in Eastern Khanty discourse. (35a) refers to a specific feast whereas (34b) does not refer to any specific referent, but is a generic referent. Additionally, the feast in example (34a) is trackable since it appears many times in the discourse (Sosa 2017: 129).

	Topical nominal object	Focal nominal object	Pronominal object
Eastern Khanty	Nom + Obj.c	Nom + Subj.c INSFIN + Subj.c	Acc + Subj.c

Table 4: DO marking in Group 3

5. Conclusions

Our analysis also gives a result of three sub-groups in the Ob-Ugrian area. These groups, however, are not identical to those in Kulonen's analysis. In our data, Northern Mansi, Northern Khanty and Southern Khanty belong to the first group, whilst the second group consists of three Mansi variants: Eastern, Western and Southern Mansi. Similar to Kulonen's model, our data also places Eastern Khanty as a sole member of the third group. We do not refer to the groups by cardinal directions, because the groups do not follow them as clearly as in Kulonen's data.

As our data show, the features of both Subject and Object Marking provide a continuum, where the morphological means used for marking the Subject or the DO diversify while moving from one group to another. The continuum is not clearly geographical: the most minimalistic system is found in the middle, the next in order is West, and the most diverse system is found in the East. The most Eastern variant, Eastern Khanty, covers more variation than any other dialect. As the division of the groups proves, the differences do not follow language boundaries but areal borders: for example, the Northern variety of Khanty is, with regard to the mentioned features, closer to Northern Mansi than to Eastern Khanty.

	Group 1	Group 2	Group 3
Verb agreement	Opposition between topical and focal DO's expressed with verb agreement.		
Case marking of nominal DO's	Canonical Subject marking Canonical DO marking	Canonical Subject marking Non-canonical DO marking: two case variants	Non-canonical Subject marking Non-canonical DO marking: three case variants
Language forms / dialects involved	Northern Mansi Northern Khanty Southern Khanty	Eastern Mansi Southern Mansi Western Mansi	Eastern Khanty

*Table 5: Areal division of Ob-Ugrian language forms according
to information structural features.*

The main principles for expressing information structure are the same in the whole Ob-Ugric community. Pragmatic variation between active and passive and between different ditransitive structures is the base of the system. The most topical core argument, the primary topic, is always placed in the most prominent syntactic role: it is realized as a Subject. With only one minor exception, the Subject is always an unmarked category, both in active and passive. In other words, the appearance of the most topical argument is connected to ZERO morphology.

With regard to DO marking, ZERO morphology is in every dialect somehow connected to focality, and case marking to topicality. Even in those variants where all the DO's appear without case marking, the topical objects are encoded in the verb. The use of the object conjugation as a marker of topicality on the DO is a common feature for all Ob-Ugrian language vari-

ants. In dialects with non-canonical marking, the topical DO's coincide with double-marking (case marking + verb agreement), whilst the focal DO's are entirely unmarked.

With both Subject Marking and DO Marking, there is diversity within the Ob-Ugrian branch. In both cases, the most variable system is found in the East corner of the area: in Eastern Khanty. All Ob-Ugrian language forms distinguish between focal and topical DO's either by verb agreement alone, or by double marking (which also includes case marking). Further, the marking of nominal and pronominal DO's is not identical in any language forms. Even if there is no case marking with nominal DO's, there is accusative marking with pronominal DO's (Northern Mansi, all Khanty variants). When the topical nominal objects are accusative marked, the pronominal DO's might be lative marked (Southern Mansi).

One interesting feature is the use of oblique cases in Subject and DO marking. As described in the previous sections, some Khanty and Mansi variants use the oblique cases (as the instructive-final in Khanty and the lative in Mansi) for marking a DO as well. The functions of these oblique cases, however, differ in Eastern Khanty and Western Mansi. In Eastern Khanty, the focal DO is marked with the oblique case, instructive-final case. In Western Mansi, the oblique case is the only case variant for marking topical DO's.

The reasons for this kind of variation are expected to stem from language contacts. As mentioned in the introduction of this article, different dialects have experienced different contacts with the surrounding languages. The influence of these contacts in detail is a question of our forthcoming studies: in this article we have presented the variation itself.

Abbreviations:

ABL	ablative	INSFIN	instructive-final
ACC	accusative	LAT	lative
DER	derivational suffix	PASS	passive
DIM	diminutive suffix	PL	plural
DO	Direct Object	POSS	possessive suffix
DU	dual	PT	present tense
FOC	focality marker	PTCL	participle
IMP	imperative	PST	past tense
INF	infinitive	SG	singular
INSTR	instrumental	TOP	topicality marker

References

Aissen, Judith 2003: Differential Object Marking: Iconicity vs. Economy. Natural Language & Linguistic Theory 21/3: 435–483.

Anderson, Gregory D. S. 2006: Towards a typology of the Siberian linguistic area. In: Matras, Yaron – Mc Mahon, April – Vincent, Nigel (eds), Linguistic Areas Convergence in Historical and Typological Perspective. Palgrave Macmillan, Hampshire. 266–300.

Bakró-Nagy, Marianne 2018: Mansi Loanword Phonology: a Historical Approach to the Typology of Repair Strategies in Russian Loanwords in Mansi. In: Bartos, Huba – den Dikken, Marcel – Bánréti, Zoltán – Váradi, Tamás (eds), Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics. Studies on Natural Language and Linguistic Theory 94: 1–66.

Bossong, Georg 1986: Differentielle Objektmarkierung in den Neuiranischen Sprachen. Gunter Narr Verlag, Tübingen.

Campbell, Lyle 2006: Areal Linguistics: A closer Scrutiny In: Matras, Yaron – Mc Mahon, April – Vincent, Nigel (eds), Linguistic Areas Convergence in Historical and Typological Perspective. Palgrave Macmillan, Hampshire. 1–31.

Croft, William 2003: Typology and Universals. Second edition. Cambridge University Press, Cambridge.

Csepregi, Márta – Sosa, Sachiko 2009: The comparable Samples of Surgut Khanty in 1996 and 2008. Journal de la Société Finno-Ougrienne 92: 193–208.

Csepregi, Márta 1998: Szurguti osztják chrestomathia. Studia uralo-altaica supplementum 6. Szeged.

Csepregi, Márta 2011: Szurguti hanti folklór szövegek. Budapesti Finnugor Füzetek 22. Budapest.

Dalrymple, Mary – Nikolaeva, Irina 2011: Objects and information structure. Cambridge University Press, Cambridge.

Filchenko, Andrey 2010: Aspect of the Grammar of Eastern Khanty. Federal Agency of Education Tomsk State Pedagogical University, Tomsk.

Filtchenko, 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. John Benjamins, Amsterdam. 47–82.

Haspelmath, Martin 2010: Comperative concepts and descriptive categories in cross-linguistic studies. Language 86: 663–687.

Honti, László 1969: A tárgy jelölése a vogul nyelv tavdai nyelvjárásában. Nyelvtudományi Közlemények 71: 113–120.

Honti, László 1984: Chrestomathia ostiacica. Tankönyvkiadó, Budapest.

Hopper, Paul – Thompson, Sandra 1980: Transitivity in Grammar and Discourse. Language 56: 251–299.

Iemmolo, Giorgio 2011: Toward a typological study of Differential Object Marking and Differential Object indexation. Dissertation. University of Pavia.

Kálmán, Béla 1961: Die russischen Lehnwörter im Wogulischen. Akadémiai Kiadó, Budapest.

Keresztes, László 1989: Mansi. In: Abondolo, Daniel (ed.), The Uralic Language. Routledge, London – New York. 387–427.

Klumpp, Gerson 2012: Differential object marking and information structure: on the function of two different pronominal accusatives in Komi and Khanty dialects. ESUKA–JEFUL 3/1: 343–372.

Kulonen, Ulla-Maija 1989: The Passive in Ob-Ugrian. Mémoires de la Société Finno-Ougrienne 203. Helsinki.

Kulonen, Ulla-Maija 1991: Über die ergativischen Konstruktionen im Ostjakischen. Journal de la Société Finno-Ougrienne 83: 181–201.

Kulonen, Ulla-Maija 2007: Itämansin kielioppi ja tekstejä. Apuneuvoja suomalais-ugrilaisten kielten opintoja varten XV. Finno-Ugrian Society, Helsinki.

LS = Лӯима сӭрипӧс [Luima Seripos]. Mansi newspaper, published twice a month in Khanty-Mansijsk.

Mark = Марк Хум Хансум Ёмас Ляххал [Gospel of Mark in Northern Mansi]. Raamatunkäännösinstituutti, Helsinki, 2000.

Moldanov, T. 2001: Молданов, Тимофей, Земля кошачьего локотка (Кань кунш оланг).Томский государственный университет, Томск.

Muysken, Pieter (ed.) 2008: From Linguistic Areas to Areal Linguistic. Studies in Language Companion Series, 90. John Benjamins, Amsterdam.

Nemysova, E. – Kajukova, L. 2007: Немысова, Евдокия – Каюкова, Людмила, Хантыйские загадки. На казымском и сургутском диалектах. Полиграфист, Ханты-Мансийск.

Nikolaeva, Irina 1999: Ostyak. Lincom Europa, München.

Nikolaeva, Irina 2014: A Grammar of Tundra Nenets. Walter de Gruyter, Berlin – Boston.

OA III = Sauer, Gert. (ed.): Ostjakologische Arbeiten 3. Akadémiai Kiadó, Budapest, 1989.

Paasonen, Heikki – Vértes, Edith 1980: H. Paasonens südostjakische Textsammlungen. Band I. Mémoires de la Société Finno-Ougrienne 172.

Pápay, József 1906: Északi-osztják nyelvtanulmányok. Nyelvtudományi Közlemények 36: 345–398.

Pápay, József 1908: Északi-osztják nyelvtanulmányok. Nyelvtudományi Közlemények 38: 313–329.

Rédei, Károly 1968: Nord-ostjakische Texte (Kazym-Dialekt) mit Skizze der Grammatik. Vandenhoeck und Ruprecht, Göttingen.

Ruttkay-Miklián, Eszter 2002: Az alany és tárgy jelölése a hanti nyelvben (különös tekintettel az ergative szerkezetre). Budapesti Uráli Műhely III. MTA Nyelvtudományi Intézet. 114–137.

Salo, Merja 2009: Hantin kielen historia ja tulevaisuus. Idäntutkimus 2009/2: 57–69. http://www.helsinki.fi/idantutkimus/arkisto/2009_2.htm

Skribnik, Elena 2001: Pragmatic structuring in Northern Mansi. In: Seilenthal, Tõnu (ed.), Congressus Nonus Internationalis Fenno-ugristarum, Pars VI: 222–239.

Solovar, V. N. 1995 = Соловар, В. Н., Модальные компоненты хантыйского предложения (тезисы). In: Абогирены Сибири: Проблемы изучения исчезающих языков и культур: Тезисы Международной научной конференции. Том 1: Филология. Институт филологи СО РАН, Новосибирск.

Sosa, Sachiko 2008: Information structure in Surgut Khanty. Paper presented in the conference Cognitive and Functional Perspective on Dynamic Tendencies in Languages on 31. May. 2008. University of Tartu.

Sosa, Sachiko 2017: Functions of morfosyntactic alternations, and information flow in Surgut Khanty discourse. Dissertation. University of Helsinki.

Steinitz, Wolfgang 1959: Zu den samojedischen Lehnwörtern im Ob-Ugrischen. Ural-Altaische Jahrbücher 31: 426–453.

Steinitz, Wolfgang 1961: Einige türkische Lehnwörter im Ostjakischen. Acta Orientalia Academiae Scientiarum Hungaricae 12: 179–181.

Toivonen, Y. H. 1945: Türkische Lehnwörter im Ostjakischen. Journal de la Société Finno-Ougrienne 52: 1–20.

Toulouze, Eva 2003: The Forest Nenets as a double language minority. Pro ethnologia, Eesti Rahva Muuseum. Multiethnic Communities in the Past and Present. 95–108.

Trubetzkoy, Nikolai Sergeevich 1928: [Proposition 16]. Acts of the First International Congress of Linguists. Leiden. 17–18.

Vajda, Edward J. 2004: The languages of Central Siberia. Introduction and overview. John Benjamins, Amsterdam.

Verbov, G. 1936: Вербов, Г., Лесные ненцы. Советская Этнография 1936/2: 56–70.

Virtanen, Susanna 2014: Pragmatic Direct Object Marking in Eastern Mansi. Linguistics 52/2: 391–413.

Virtanen, Susanna 2015: Information Structure in Eastern Mansi. An Information Structural Approach. Dissertation. University of Helsinki.

WV = Kannisto, Artturi 1951–1963 (toim. Matti Liimola): Wogulische Volksdichtung I–VI. SUST 101, 109, 111, 114, 116, 134. Helsinki.

Online sources:

Northern Mansi corpus / Szilágyi:
http://norbertszilagyi91.wixsite.com/tawdamansi/corpora

Western Mansi and Northern Khanty corpus:
http://www.babel.gwi.uni-muenchen.de