Uralic studies, languages, and researchers

Edited by Sándor Szeverényi

Studia uralo-altaica 54

Uralic studies, languages, and researchers

Proceedings of the 5th Mikola Conference
19–20, September 2019

Edited by Sándor Szeverényi

Szeged, 2021

Table of contents

A word-formational approach to neologisms in modern Northern Mansi: Bilingual compound names of professions, documents and institutions

Susanna Virtanen

Helsinki University

1. Introduction

Neologism are new words recently adapted to the language. Usually neologisms are created by borrowing from other languages or by deriving and compounding already existing words, or by bringing up wholly new coinages. This article is one part of a more extensive study on Mansi neologism. In this article, I will outline one word-formational group of neologisms in modern Northern Mansi: bilingual (Mansi-Russian) compounds. I will concentrate especially on three thematic groups: names of professions, official documents and institutions. These thematic groups were chosen, because they also form the most frequent word-formational types of bilingual compounds appearing in my data. In other words, there is a connection between thematic groups and word-formational types.

Mansi language belongs to the Ob-Ugric branch of the Finno-Ugrian language family. It is spoken in Western Siberia, in the Khanty-Mansi Autonomous Area of the Russian Federation. Mansi is an urgently endangered language: there are still approx. 1,000 speakers alive (see the Russian Census 2010). Only one of the four main dialects is still spoken. The speakers of Mansi live under the Russian laws and inside the Russian society. In elementary schools, Mansi is often taught as a second language only, and it is not used as a teaching language at all. Mansi is not used in everyday situations by authorities or public offices. Urbanization and oil drilling speed up the death of the language. For this reason, there are also a lot of Russian loanwords in Mansi language.

The data for my study is gathered from up-to-date newspaper articles. Methodologically this study combines together word-formational analysis and semantics: the results of word-formational and thematic classification processes are compared with each other. The aim of this study is to present a formal description of

bilingual compounds (Russian-Mansi combinations), and to describe, what kind of connections or correlations there are between word-formational types and thematic groups. Bilingual compounds include either a Russian generic part and a Mansi specific part, or a Mansi generic part and Russian specific part. The semantic properties of both types will be observed.

2. Language contacts of Mansi

During the centuries, Mansi has been under influence of several surrounding languages. Proto-Iranian and Proto-Turkic influence started already in the common Ugric era and continued after Hungarian broke away (Keresztes 1989: 422). Among the Uralic languages, Komi, Khanty and Nenets have influenced Mansi. Contacts with Komi people started in the 900's (Keresztes 1989: 423), and also the first Russian loanwords were adapted with Komi mediation. Also Samoyedic influence has been observed: especially the Mansi reindeer-breeding terminology has been influenced by the Nenets language (Steinitz 1959). Southern dialects of Mansi have also been in contact with Tatar languages from the 14th Century (Keresztes 1989: 422). With the Russian population, there have been direct contacts since the 17th Century (see e.g. Kálmán 1961, Bakró-Nagy 2018).

Later on, a great deal of new words have been adapted to Mansi from Russian when adapting the whole phenomenon from the Russians. This covers a large amount of political, technological, scientific, cultural and administrational vocabulary, and vocabulary on other thematic groups as well. The current sociolinguistic situation is that Mansi is used only in a restricted area of everyday life: many situations are excluded from Mansi language using area, because for example administration and education are organized mainly in Russian. Practically all speakers of Mansi are Russian-Mansi bilinguals, and for a big majority of them, Russian is the language they speak best. This means that Russian language strongly influences Mansi language and makes maintaining the minority language much more difficult. One of the aims of this study is to map out the circumstances where new phenomena are named with Mansi-origin words are used in neologisms.

3. Research on neologisms

As for example Häkkinen (1997: 85–87) states, new words come to a language first and foremost by compounding and deriving old words, by borrowing words from other languages and sometimes by inventing completely new coinages. Neologisms are generally defined as words that have recently come into the language or have not been established yet. For example, Bolganbaev (1988: 112) defines neologism as follows: "Neologisms are words that have appeared in a language in connection with

new phenomena, new concepts… but which have not yet entered into the active vocabularies of a significant portion of the native speakers of the language." Fischer's (1998: 3) definition is basing on similar grounds: "A neologism is a word which has lost its status of a nonce-formation but is still one which is considered new by the majority of the members of a speech community."

The definitions above emphasize that neologisms are connected to new cultural concepts and phenomena, and also that neologisms are not yet established in the language. The connection to new cultural concepts is the starting point of gathering my own data as well: my data focuses on the modern society and culture of the 20th and 21st centuries. This kind of cultural features have come to the Mansi community almost exclusively through Russian society. In this study, such lexical items are presented, which are connected to modern society and culture: technology, health/medicine, cultural life, modern professions and administration/society.

Similar to other types of lexicology, also neologisms can be classified according to their functions, coinage processes, formation, etymological sources, or structural and organizational feature. Further, a formational classification includes several categories. Smyk-Bhattacharjee (2009: 36), referring to Helfrich (1993) classifies neologisms into four groups:

1. morphological neologisms
2. semantic neologisms
3. loan neologisms
4. coinage neologisms

My classification is quite similar to that of Smyk-Bhattacharjee, but my terminology and way of classification differs slightly from hers: I have ended up using partly more punctual terms *derivation* (morphological neologism), *metaphor* (semantic neologism), *loanword* and *compound*. Shortenings or coinages are not included in this study.

This article is one part of a larger lexical study on neologisms and loanwords in modern Northern Mansi. The more extensive study is a descriptive study on the formation and semantics of Mansi neologisms, involving mainly formational perspective. The study includes three sub-studies: Russian loanwords, Mansi-origin words and bilingual compounds (see Figure 1 in Section 5.1). As described in detail in Section 4, this partial study is a more detailed approach to one word-formational part of my data.

4. Research data

My data are gathered from the volumes 2014, 2018 and 2019 of *Lūimaa Sēripos*, the only regularly published Mansi-language newspaper in the world. It is published twice

a month in Khanty-Mansijsk and is also readable online[1]. The data have been gathered in a traditional way by reading texts and picking up words belonging to particular thematic groups: all groups are connected to technological, political and cultural innovations of the 20th and 21st centuries. The data have been classified according to two criteria: word-formational category and thematic category. The whole data include a corpus of 500 lexical items, including the following thematic groups:

1. Modern professions
2. Technology, Computing, Internet
3. Modern Society, Administration
4. Education, Cultural life, Science
5. Health care, Medicine

Further, the data have been classified into Russian loanwords, Mansi-origin words and bilingual compounds (see Figure 1 in Section 5.1). The need for this kind a classification originated from the very nature of the data: the details were decided just after observing the content of the data first.

In this case, the demarcation of the data is not straightforward, and the age of a single word cannot always be defined exactly, especially in case of a less documented language. The starting point of this study is to collect words belonging to thematic groups that have only emerged in human culture during the last hundred years. At this stage (2021), not all of the words in my data necessarily meet the definitions above, but the starting point is that the words have at least been neologisms once within the past 100 years.

5. Neologisms in my data

In this section, the above described data is first presented as a whole in 5.1. Further, the different bilingual compound types are presented in 5.2 and 5.3.

5.1 Overview of my data

The data for this study have been classified by two different perspectives: the thematic and the word-formational one. The thematic groups were already presented in Section 4. This article concentrates only on one part of the data, but first it is appropriate to present the ford-formational classification of my whole data. The classification is presented in Figure 1:

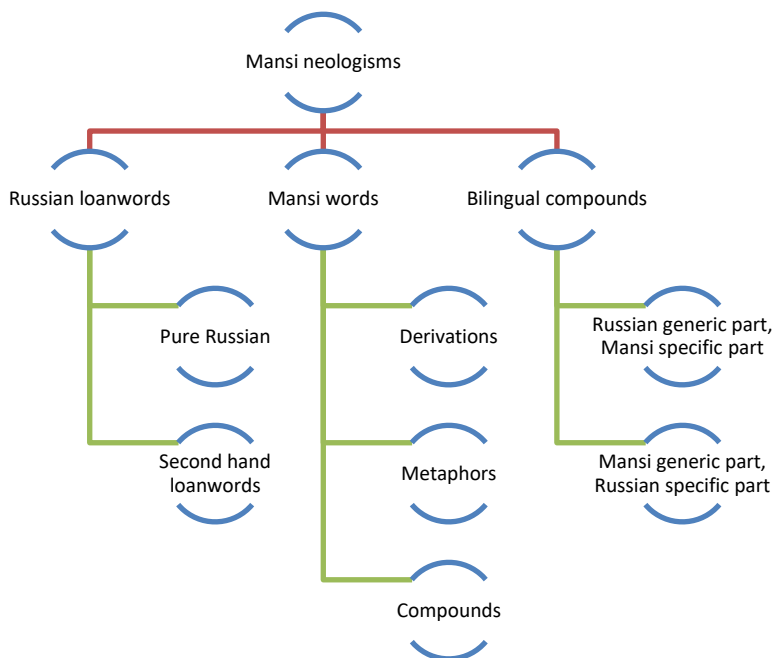[1] http://www.khanty-yasang.ru/luima-seripos

Figure 1. Word-formational classification of my data.

This study is an analysis of one part of the above mentioned data: I will concentrate on bilingual compounds (the right periphery of Figure 1). Bilingual compounds make up 39% of the whole data. They are mostly used in profession names, institution names, document names, other bureaucratical terms and medical terminology. I have divided the bilingual compounds into two types: the 'Russian specific part + Mansi generic part' type (see Section 5.2) appears much more often than the 'Mansi specific part + Russian general part' type (see Section 5.3). The first mentioned type can also be further divided into thematic sub-types.

The two types differ from each other both word-formationally and semantically. First, they differ in thematic groups connected to them. Secondly, there are clear differences between these two types especially in the way the generic parts behave in compounds. Both types are, however, productive. In the following subsections, I will present the most productive types and sub-types of bilingual compound in my data from both thematic and word-formational perspective. The similarities and differences between the two main types of compounds will be described in detail. Compounds with a Mansi generic part are presented in Section 5.2, and those with a Russian generic part in 5.3.

5.2. Russian specific part, Mansi general parts

In the following subsections 5.2.1–5.2.3 I will present three thematic groups (sub-types of one of the main types) that are connected to the 'Russian specific part + Mansi generic part' types compounds. These are compound document names (5.2.1), compound institution and building names (5.2.2) and compound profession names (5.2.3). The subsections are named by the thematic groups, and under these titles, also the word-formational properties are observed in detail.

5.2.1 Compound document names with a Russian specific part and the Mansi generic part *нэпак* 'book; letter'

One of the most productive bilingual compound types in my data are the document names created with a Russian specific part and the Mansi generic part *нэпак* 'book; letter'. The Russian specific part is always a Russian document name, and it remains unmodified.

All of the examples in this group are created in the same way, for example: *документ-нэпак* 'document', *полис-нэпак* 'insurance document', *виза-нэпак* 'visa', *диплом-нэпак* 'diploma, examination certificate', *договор-нэпак* 'contract', *заявка-нэпак* 'application', *заявление-нэпак* 'application', *квитанции-нэпак* 'receipt', *реквизит-нэпак* 'personal information form'.

Also some book names are created in the same way: *букварь-нэпак* 'ABC Book', *диссертация-нэпак* 'PhD Thesis', *атлас-нэпак* 'atlas'. Still, the formation methodology is exactly the same: a Russian term is adapted to Mansi language by turning it into a compound and adding a Mansi word as a generic part. Further, just as a side note and for comparison, also more typical bilingual compounds are found in the data, like *программа-нэпак* 'program booklet', *сāккон-нэпак* 'statute book' (Russ. *закон* > *сāккон*), where the specific part is specifying the generic part in a normal way.

A common feature for all examples in this group is that they include a Russian document (or book) name, which is turned into the specific part of a compound by adding the Mansi term *нэпак*. *Нэпак* serves as a generic part: from this point of view the specific part's function is to tell, what kind of document is in question. However, the meaning of the original Russian document name remains the meaning of the compound as well: semantically it already includes the meaning 'document'. In other words, the function of the Mansi generic part is more likely to adapt the term to Mansi language. The Mansi generic part gives us the information that there is talk about a document, while the same information is included in the specific part as well.

5.2.2 Compound names of institutions and buildings, with the Mansi generic part *кол* 'house'

In the same way as 'нэ̄пак' in 5.2.1, also the Mansi word *кол* 'house' functions as a generic part of bilingual compounds: compounds produced with *кол* are names of institutions or concrete buildings. Names in this group include a Russian institution name or a name of a building as a specific part, and the Mansi term *кол* as a generic part. Further, both appellatives and proper names of institutions are included as specific part.

Appellatives appearing as specific parts typically refer to institutions maintained by society: *садик-кол* 'kindergarten' *библиотека-кол* 'library', *интернат-кол* 'dormitory', *театр-кол* 'theatre', *клуб-кол* 'club', *консерватория-кол* 'conservatory' *архив-кол* 'archive' *профсоюзов кол* 'trade union office' *радио кол* 'radio house', *спортивный яныг кол* 'sport center',

In addition to appellatives presented above, also many proper names are turned into compounds in the same way: *Соссаӈ мир культура кол* 'Indigenouse Peoples' Cultural Center', *Центр искусств кол* 'Art Centre', «*Институт развития образования*» *кол* 'Center for Development of Education', *Центр профессиональной патологии кол* 'Center for Occupational Pathology'. All of these are Russian proper names just refined with the Mansi word *кол*.

Again, the meaning of the Russian generic part remains the meaning of the whole compound. Technically the Mansi part is a generic part, and it includes the information that there is talk about a building or an institution, but the same meaning is included already in the specific part. Again, the Mansi generic part functions as an "adapter" to Mansi language.

5.2.3 Compound profession names with Mansi generic part *хо̄тпа* 'person', *о̄йка* 'old man', *ма̄хум* 'people', *хум* 'man' or *нэ̄* 'woman'

Third similar group to 5.2.1 and 5.2.2 are the compound profession names. They are produced with a Russian profession name as a specific part, and a Mansi human-related word as a generic part. Typical generic parts are *хо̄тпа* 'person', *о̄йка* 'old man', *ма̄хум* 'people', *хум* 'man' and *нэ̄* 'woman'. In other words, the generic part varies depending on the gender and the age of the professional. Naturally, there may be different variants of one profession name according to gender and age.

Some of the names are written with a dash, some not: there seems not to be any consistent practice for using or not using the dash. For example: *юрист-хо̄тпа*

'lawyer', *учёный ōйка* 'male researcher', *жюри хōтпа* 'jury member', *яныг лӭккар-хōтпа* 'senior physician' ['a big doctor-person'], *лӭккар-мāхум* 'medical staff'.

This thematic group includes also compounds, where the specific part consists of a Russian noun and a participle form of a Mansi verb: Russian *кина вāрнэ мāхум* 'film industry people' ['people making films'].

5.3. Mansi specific part, Russian generic part

Compound words with a Mansi specific part and a Russian generic part differ structurally from the examples presented in 5.2.1–5.2.3. Further, they do not form as consistent and homogenous group as the three types in the previous subsections. A common feature for all examples also in this group is that the Russian generic parts of the compounds represent phenomena of modern society, adapted from the Russian culture.

The Russian generic parts are specified by Mansi terms. For example *сым āгм пусмалтан пӯльница* 'hospital for heart diseases', *мā-вōй нōх-винэ компаниян* 'oil drilling company', *щāнь лāтӈут ӯргалан Фонд* 'fund for protecting minority languages', *ханищтан программа* 'teaching programme', *ханищтап департамент* 'department of education', *сым пусмалтан лӭккар* 'oculist'.

The generic parts of the compounds presented in this subsection represent cultural phenomena adapted from Russians to Mansi society. Differently from the examples presented in the previous subsections, the roles of specific and generic parts of the compounds' follow the normal semantic structure of compounds. The Mansi specific part is really specifying the generic one: Russian general terms are specified by Mansi expressions. Old words in both old and new meanings are used.

5.4. The most productive compound types in brief:

To sum up, the Russian profession, institution or document names build up a special type of compounds (see 5.2.1–5.2.3). Russian names of professions, institutions and documents are turned into specific parts of compounds. All of them are added a Mansi generic part, which is semantically a quite simple and concrete phenomenon. The Russian terms are added a semantically simple Mansi word (generic part) to produce a compound, while the meaning of the compound remains the same as the original term's meaning in Russian. The function of the generic part is not to bring new information on the target but to adapt the term into Mansi language.

Is this a special word-formation type or a simple compound? Technically it is a very typical compound: the generic part expresses the general nature of the phenomenon, and the specific part gives more specific information on it. However, word-formationally and semantically the question is more complicated: the specific

part already includes the same information as the generic part. Also Hauel (1993) has observed similar lexical compounds in Khanty, when investigation Russian-origin loan translations in modern Khanty. Hauel's perspective and starting point are different from those in this study, but her data include examples similar to those in my data.

Hauel mentions Khanty profession names very similar to the Mansi examples presented above: also in Khanty, profession names are created with a human-related word like *нэ* 'woman', *хоят* person', *хе* 'man' and *ёт* 'people' as a generic part (Hauel 1993: 122). Also words *вер* 'work', *ут* 'thing' and *сир* 'manner' are used in compounds in the same way. Differently from the Mansi data presented here, they usually include a Khanty-origin specific part, or a Russian-origin word adapted to Khanty voice system. Hauel compares the appearance of loan translations (Lehnübersetzung), loan renditions (Lehnübertragung) and loan creations (Lehnschöpfung) (Hauel 1993: 123). All of these categories are not comparable with Mansi bilingual compounds, especially not with those with a Russian specific part and a Mansi generic part. In other words, many of them are nearer to the Mansi-origin derivations or compounds, which are not discussed in this partial study (but many appear in my data). However, the Khanty generic parts and their nature are interesting from the point of you of my study as well.

Hauel calls these generic parts of compounds 'half-suffixes' (Hauel 1993: 123). This refers to the generic parts' derivational nature: according to this point of view, they are not clearly parts of compounds but more likely derivational suffixes. The same idea can be applied to Mansi data as well: from one perspective, the Russian terms are adapted to Mansi language by a special way of derivation, by creating compounds, which actually are Mansi derivations of Russian-origin words.

It is good to remember that beside bilingual compounds, also monolingual Russian or Mansi synonyms are used. If we turn back to Figure 1, some phenomena may have equivalents in every branch of the figure, or at least in two of them. For example all three variants *мāнь нÿврамыт кол*, *садик* and *садик-кол* appear in the meaning of 'kindergarten'. Both *договор* and *договор-нӭпак* appear for 'contract'. Basing on this information, we can say that providing a derivation-like bilingual compound is not obligatory from any perspective. It is one way to bring new words to language, and get them better adapted, but not the only one.

Further, these data were gathered from newspapers. Newspaper language is more official and more formal than some other registers of communication. It is possible that some different kind of data would show different results. By investigating newspaper language we cannot find out, whether people use loanwords or bilingual compounds or something else for example in oral discussions or private letters.

6. Conclusions and further questions

The data presented in this study show that bilingual compounds appear especially in the field of society, professions and bureaucracy. The productive types of bilingual compounds in my data are:

1) "Mansi derivations" of Russian neologisms (69%), including document names with a Russian specific part and the Mansi generic part *нэпак*, institution names with a Russian specific part and the Mansi generic part *кол* and profession names with a Russian specific part and a Mansi generic part referring to human beings.

2) Institution and organization names with a Mansi specific part and a Russian generic part (21%).

3) Other types 10%.[2]

A common feature for all is that they are semantically related to phenomena adapted from the Russian society. Types 1 and 2 differ from each other word-formationally. In type 1, the function of a Mansi generic part is to adapt the Russian term into Mansi language: the meaning of the specific part becomes the meaning of the new compound as well. In type 2, a Russian generic part is specified by a Mansi specific part in a normal way.

Finally, to present some further questions and perspectives, the topic could be further elaborated by qualitative, quantitative or typological aspect. A qualitative approach includes the question of productivity of the compound types with a Mansi generic part like *нэпак*, *кол* or *хōтпа*: which other generic parts can this model be extended to? Also a quantitative approach on the topic will be useful: with how many specific parts do these generic parts appear?

Finally, a typological survey by comparing the Mansi data with other languages, also those with a different sociolinguistic status, would give many answers. Are there parallel situations in other languages of the world? Is this a feature of an endangered minority language only? Is this a feature of a typologically particular language? Do compounds like this appear in any other sociolinguistic contexts? These questions are topics of further studies, and investigating them will perhaps enlighten the motivation for creating untypical compounds in Mansi as well.

---

[2] These types or cases are not discussed here, because they are quantitatively marginal cases.

References

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 and Váradi, Tamás (eds.), *Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics.* Studies on Natural Language and Linguistic Theory 94.

Bolganbaev, A. 1988. Lexicology in Kazakh language. *Almaty Mektep.* 14–23.

Fischer, A. 1989. Aspects of Historical Lexicology. In: Fries, U., and Heusser, M. (eds.), *Meaning and Beyond: Ernst Leisi zum 70 Geburtstag*. Tübingen: Gunter Narr. 71–91.

Hauel, Petra 1993. Ostjakische Lehnübersetzung im Bereich der Nominalen Zusammensetzung. In: Bakró-Nagy, Marianne and Szíj, Enikő (eds.), *Hajdú Péter 70 éves*. Linguistica Series A. Studia et Dissertationes 15. Budapest: MTA Nyelvtudományi Intézet.

Helfrich, U. 1993. *Neologismen auf dem Prüfstand. Ein Modell zur Ermittlung der Akzeptanz französischer Neologismen.* Wilhelmsfeld: Gottfried Egert.

Häkkinen, Kaisa 1997. *Mistä sanat tulevat?* Tietolipas 117. Second edition. Helsinki: Finnish Literature Society.

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

Keresztes, László 1998. Mansi. In: Daniel Abondolo (ed.), *The Uralic languages*. London/New York: Routledge. 387–427.

Smyk-Bhattacharjee, Dorota 2009. *Lexical Innovation on the Internet – Neologisms in Blogs*. Doctoral Thesis. Faculty of Arts, University of Zurich. Steinitz, Wolfgang 1959. Zu den samojedischen Lehnwörtern im Ob-Ugrischen. *Ural-Altaische Jahrbücher* 31: 426–453.