As a manuscript

Anna Yu. Urmanchieva

Reconstruction of the linguistic landscape of Western Siberia
(based on Samoyedic languages)

Dissertation Summary
for the purpose of obtaining
academic degree Doctor of Science in Philology and Linguistics

St. Petersburg
2023
The work was carried out at the Department of the Theory of Grammar of the Institute of Linguistic Studies of the Russian Academy of Sciences
GENERAL DESCRIPTION OF THE WORK

The dissertation is devoted to the reconstruction of the linguistic history of Western Siberia on the basis of the Samoyedic languages. As is well known, documentary information concerning the peoples inhabiting Western Siberia and their languages actually appears only with the entry of these territories into the Russian Empire. This process dates back to the 16th–17th centuries. But the materials of the scientific study of the Samoyedic languages and the idioms adjacent to them appear a couple of centuries later, starting from the 18th century, and are associated with the names of D. G. Messerschmidt (see, for example, [Messerschmidt 1962–1977] or [Messerschmidt 2021]), G. F. Miller (see, for example, such important publications as [Helimsky 1986], [Dulzon 1961], [Elert (ed.). 1996], [Helimski 1987]), P. S. Pallas (see, for example, [Pallas 1776], [Helimski 1987]), F. I. von Strahlenberg ([Strahlenberg 1730]) and others. However, in the full sense scientific fixation of the material of the Samoyedic languages began only in the middle of 19th century with the works of M. A. Castrén ([Castrén 1854; 1855]). For example, the book [Helimski 1997] devoted to the Mator language is a striking example of the fact that vocabularies compiled even in the early 19th century, requires painstaking (and not always possible) linguistic interpretation.

That means, that historical information about the Samoyedic peoples and their languages allows us to more or less reliably imagine only the last three centuries of their history. In this situation, the methods of linguistic reconstruction acquire paramount importance, primarily the methods of comparative historical linguistics, but also the methods of linguistic contactology. The application of this set of methods to the material of the Samoyedic (and, more broadly, to the Uralic languages) has so far made it possible to localize the linguistic homeland of the Samoyeds. In [Helimsky 1989/2000 and 1996/2000], it is localized “in the Chulym-Middle Yenisei region, approximately around the triangle “Tomsk-Krasnoyarsk-Yeniseisk”; this point of view is also supported in [Napolskikh 2002/2018]. Naturally, there are other hypothesis concerning the localization of the Samoyedic homeland. For example, J. Janhunen places it in the Minusinsk Basin [Janhunen 2022]. But within the current study the localization of homeland proposed by E. A. Helimsky is accepted.

At the same time, in the case of the Samoyedic languages, we have a significantly worse idea of the history of individual branches of the Samoyedic group and the history of individual languages — the descendants of the Proto-Samoyedic. In the historical period, we find the Kamas-Koibal (hereinafter Kamas) and Mator-Taigi-Karagas (hereinafter Mator) in the foothills of the Sayan Mountains; the ancestors of peoples speaking Tundra Nenets, Tundra and Forest Enets and
Nganasan entered the Arctic tundra zones; Selkups and Forest Nenets remained in the taiga zone of the Yenisei-Ob interfluve. But what were the migration processes, that led the ancestors of the speakers of these Samoyed languages to the modern territories of their settlement (for the sake of simplicity, I apply the term “modern” to both Kamas and Mator, too, meaning the localization of these two now dead languages in the historical period)?

Thus, it turns out that we have a more or less substantiated localization of the Samoyedic homeland, as well as historical information about the territories inhabited by speakers of modern Samoyedic languages, but we know almost nothing about the migration processes that took place in that more than one and a half thousand years period that separates the era of the Samoyedic homeland from the era of modern Samoyedic languages.

In this study, an attempt was made to fill this gap with linguistic methods. First of all, the external relations of the Samoyedic languages with their potential linguistic neighbors were studied. However, it should be noted that when studying external relations, I do not dwell in detail on the ancient connections of the Samoyedic languages with the languages of other linguistic families — Indo-European, Turkic, Mongolian, Manchu-Tungusic (the only exception is the study of the relative chronology of the Yenisei-Selkup contacts): such studies have already a solid history. I have focused on the problem of Samoyedic=Ob-Ugric linguistic contacts: as far as these languages belong to the same Uralic family, a strong tendency exist to regard all parallels found between the two branches as representing a common Uralic protoform — but among these parallels one can also find those that should be considered as common areal innovations.

The main purpose of Samoyedic=Ob-Ugric language contacts study is the reconstruction of the mutual position of the Proto-Samoyedic and Proto-Ob-Ugric languages. This study allows to answer two questions. The first one: Did the contacts of the Proto-Ob-Ugric and Proto-Samoyedic languages continue after the divergence of Proto-Uralic? The second one: which of the Ob-Ugric and Samoyedic languages (after the divergence of the two protolanguages) continued to maintain mutual influence?

After carrying the study of external relations of the Samoyedic languages within the Uralic language family, it is necessary to turn to the study of internal relations within the Samoyedic group of languages. First of all, it is necessary to revise the classification of the Samoyedic languages. At the moment, there are at least five of them classifications (and that is obviously too much, considering that only six elements are classified). And in general, all the classifications in question offer quite different models of the divergence of Samoyedic languages. And it is obvious that a correct classification scheme will bring us significantly closer to the reconstruction of the historically correct mutual localization of languages on the map and to the reconstruction of the migration routes that led to that mutual localization.
An equally important issue is the reconstruction of contact situations that, in different historical periods, connected the languages within the Samoyedic group. Obviously, a set of such contact situations must also correspond to several spatial models of the mutual localization of the Samoyedic languages in different historical periods. Such areal convergence models, together with genetic divergence models, will make it possible to more reliably reconstruct most significant changes of the linguistic map of Western Siberia, starting from the period of the Proto-Samoyedic first divergence and ending with the period of the modern localization of the descendant languages.

It should also be noted that the parallel study of genetic divergence and areal convergence made it possible to clearly distinguish between linguistic phenomena that reflect connections of the both types - and, therefore, we obtain a more objective genetic classification of the Samoyedic languages, abstracting from the distortion of the picture caused by secondary convergent processes.

Thus, the current relevance of the study is determined by the poor knowledge of the intermediate stages in the history of the Samoyedic language group: by now, the Samoyedic homeland has been reliably reconstructed, but we know practically nothing about the migration processes, that led from that homeland to the areas, inhabited by the speakers of modern Samoyedic languages. Meanwhile, the linguistic reconstruction of the history of the Samoyedic language group is of great importance not only within the framework of Uralic linguistics, but also in related disciplines, primarily for the reconstruction of the ethnic history of Western Siberia.

The novelty of the study is determined by the developed comprehensive approach, in which traditional comparative study of genetic relationships within the Samoyedic group was supplemented by the reconstruction of historical contact situations within the group in different periods of its history. The results obtained are also characterized by novelty: a) lists of lexical correspondences of Nganasan with the South Samoyedic languages and Nganasan with Enets have been collected; b) the linguistic map of Western Siberia was reconstructed in dynamics; c) on the basis of the study of narrative strategies, it was established the existence of a relatively recent contact zone that united the Nganasan language, the Northern and Tym dialects of the Selkup and the Eastern subdialects of the Vakh Khanty.

The theoretical significance of the study is determined by the development of a comprehensive methodology of reconstruction of the history of a language family / language group. The following results obtained are also of theoretical significance: The following conclusions are also of theoretical significance: the fact of contacts between the Ob-Ugric and Samoyedic languages was proved, the fact of separate contacts between the Mansi and South Samoyedic languages was proved. A number of important results have been obtained regarding
the historical location of the Ob-Ugric languages on the linguistic map, and regarding divergent and convergent processes within the Samoyedic linguistic group.

The practical significance of the study lies in the possibility of using its results and materials in university courses in Uralic linguistics and Samoyedology, as well as in courses in comparative historical linguistics and linguistic contactology.

The object of the study is the Uralic languages of Western Siberia (primarily Samoyedic, but also Ob-Ugric) from the point of view of reconstructing the linguistic history of the region.

The subject of the study is the areal relationships between the Ob-Ugric and Samoyedic languages before and after the divergence of both proto-languages, genetic relationships within the Samoyedic group, as well as areal relationships within the Samoyedic language group and their historical changes.

The aim of the study is to reconstruct the linguistic map of Western Siberia in dynamics. Achieving of the formulated goal involves solving the following tasks:

1. Identification of lexical parallels, which are evidence of historical contacts between Proto-Samoyedic and Proto-Ob-Ugric;

2. Identification of lexical parallels, which are evidence of historical separate contacts between Mansi and South Samoyedic (to a lesser extent — Kamas, to a greater extent — Selkup);

3. Analysis of the geographical distribution of Turkisms in the Mansi dialect continuum and in the Khanty dialect continuum to determine the earlier location of Proto-Khanty and Proto-Mansi on the linguistic map of Western Siberia. This task follows from the previous one: such an analysis is necessary to explain why I establish early historical contacts not between Khanty and Selkup (which would be the most natural conclusion, if one takes into account the modern linguistic map) but between Mansi and Selkup (although on the modern linguistic map Mansi and Selkup are not adjacent to each other);

4. Collection of two lists of the most stable (“basic”) lexical items (M. Swadesh's 100-word list and Leipzig-Jakarta 100-word list) and their lexicostatistical analysis in order to measure the relative genetic distance between the Samoyedic languages and to identify of the most undoubted “nodes” of genetic classification. At the same time, the analysis of lexicostatistical data identified a number of additional tasks (see 5–7 below)/ Thus, for several languages it was necessary to supplement the analysis of two short lists by the analysis of the whole vocabulary available in order to find exclusive lexical parallels, connecting the language in question with the other Samoyedic languages. Further the established exclusive parallels were classified as common archaisms, common genetic innovations, and common areal innovations.

5. Clarification of the place of Mator in the genetic classification of the Samoyedic languages. The Mator vocabulary of Samoyedic origin from [Helimski 1997] was analyzed in
order to identify the corpus of exclusive Mator — Southern Samoyedic parallels and the corpus of exclusive Mator — Northern Samoyed parallels. Among these parallels innovations were identified and further analyzed in order to determine their areal or genetic nature.

6. Explanation of the relatively high percentage of lexical correspondences of Nganasan with the languages of the Southern Samoyedic branch (Selkup and Kamas). Nganasan, Selkup and Kamas lexical materials were analyzed throughout in order to identify the set of exclusive Nganasan-Selkup-Kamas, Nganasan-Selkup and Nganasan-Kamas lexical parallel. Further, among these parallels innovations were identified and analyzed in order to determine their areal or genetic nature.

7. Explanation of the relatively high percentage of lexical correspondences of Nganasan with both Enets languages (that is, with Proto-Enets). A set of exclusive Nganasan-Enets lexical parallels was identified (since, according to lexicostatistical data, Enets is undoubtedly closer to Nenets than to Nganasan, it was unnecessary to additionally show that the Nganasan-Enets parallels are of secondary, contact origin);

8. Drawing up a series of maps (based on the results of solving the research problems formulated in 1–7 above), reflecting the changes of the linguistic landscape of Western Siberia in the period starting with the divergence of the Proto-Samoyedic and ending with the settlement of speakers of Samoyedic languages in their modern territories;

9. Analysis of the repertoire of narrative strategies of the Nganasan language, Northern and Tym dialects of the Selkup and Eastern sub-dialects of the Vakh Khanty. The results of the study allow to state that in the recent historical perspective there was a contact zone that united these idioms;

The study was based on existing materials on the lexical reconstruction of the Samoyedic and Uralic languages [SW; Etym 1; Etym 2; UEW], lexical materials on Samoyedic languages: [Helimski 1997] — for Mator, [Alatalo 2004] (sometimes with additions according to [Kazakevich, Budyanskaya 2010] and [Bykonya et al. 2005]) — for Selkup, [Donner, Joki 1944], sometimes with additions according to [Castrén 1855] — for Kamas, unpublished dictionaries collected by E. A. Helimsky — for Nganasan and Enets (for Nganasan also additionally the dictionary [Kosterkina, Momde, Zhdanova 2001], for Forest Enets — the dictionary [Sorokina, Bolina 2009]), [Lehtisalo 1947] — for Tundra and Forest Nenets, for Tundra Nenets also [Tereshchenko 1965], the phonological form of Tundra Nenets words is listed according to [Salminen 1998]. Ob-Ugric lexical materials are cited: for Mansi — according to [Kálmán, Munkácsi 1985], for Khanty — according to [Karjalainen 1948]. The material for the study of narrative strategies was collected as follows: for Nganasan — texts from the collection of Nganasan texts by V. Yu. Gusev and M. M. Brykina and texts from [Katzschmann 2008] were
analyzed, for Northern and Tym Selkup — texts from the corpus of the Selkup language by J. Alatalo, for the Eastern Vakh Khanty language — texts from the archive of N. I. Tereshkin in the Institute of Linguistic Studies of the Russian Academy of Sciences and texts from [Tereshkin 1961].

The dissertation research uses the methods of comparative historical and comparative linguistics, as well as the methods of areal and contact linguistics.

The dissertation research formulates a hypothesis concerning the ethnohistorical implications of the conducted linguistic research. Complementing the comparative historical study of the Samoyedic languages with the methods of contact and areal linguistics makes it possible to reconstruct a number of sequentially replacing each other linguistic landscapes of Western Siberia.

The following are put forward for defense:

1. The contacts between Proto-Samoyedic and Proto-Ob-Ugric were not interrupted after the separation of these two proto-languages from the Proto-Uralic;

2. After the divergence of Proto-Ob-Ugric and Proto-Samoyed proto-language, separate contacts of the Mansi and South Samoyedic languages were preserved (Kamas was involved to a lesser extent, Selkup — to a greater extent); these are the most ancient separate contacts between individual Ob-Ugric and individual Samoyedic languages;

3. The earliest mutual arrangement of the Proto-Khanty and Proto-Mansi differed considerably from the modern situation. One should suggests rather a division along the “north—south” line: the Proto-Mansi occupied the more southern part of the Ob-Ugric area (and its eastern border reached the upper reaches of the Vasyugan), and the Proto-Khanty was located north of the Mansi, and it was not distributed east of the Lower Ob basin.

4. Mator belongs to the group of Northern Samoyedic languages, where it forms a single subgroup with Nenets. The exclusive lexical correspondences of Mator and Southern Samoyedic (primarily of Mator and Kamas) can be interpreted in most cases as secondary areal innovations.

5. Nganasan demonstrates a significant number of secondary contact isoglosses with Kamas and Selkup.

6. Nganasan demonstrates a significant number of secondary contact isoglosses with Proto-Enets.

7. In the recent historical perspective, there was a contact zone that united the Nganasan language, the Northern and Tym dialects of the Selkup and the Eastern sub-dialects of the Vakh Khanty.

The main results of the study were discussed at the following international and Russian scientific conferences: “International Conference on Samoyedology” (6th — Moscow, 2016, 7th — Tartu, 2018), “Conference of the European Linguistic Society” (Tallinn, 2018), “Conference
on Uralic, Altaic and Paleoasiatic languages” (St. Petersburg, 1st — 2018, 2nd — 2020, 3rd —
2022), “Languages of the minor peoples of Russia: oral vs. written” (St. Petersburg, 2017),
“Languages of the peoples of Siberia and adjacent regions” (Novosibirsk, 2019), “Language
contacts in the circumpolar area” (Moscow, 1st — 2017, 2nd — 2019), “5-th conference in
memory of T. Mikola” (Szeged, 2019), “Documentation of languages and dialects of indigenous
peoples of Russia” (St. Petersburg, 2019), “Grammatical categories in world languages: hierarchy
and interaction” (St. Petersburg, 2016). The following works have been published on the topic of
the dissertation:

Papers in journals, indexed in WoS, Scopus, or included in the list of high-ranked
journals recommended by the National Research University Higher School of Economics
767.
semantiki i pragmatiki v opisanii glagol'nykh grammem // Voprosy yazykoznaniya. 2014.
№ 4. S. 66–86.
enklitikoy nyo'' ~ nyu’’: vzaimodeystviye znacheniy // Acta Linguistica Petropolitana.
5. V. A. Plungian, A. Yu. Urmanchieva. Kontrastivnoye proshedsheye v enetskom yazyke:
unikal'nyy ''plyuskvamperfekt'' i yego ekspansiya v glagol'noy sisteme // Acta Linguistica
468.
6. A. Yu. Urmanchieva. Ot imperfektivnosti k evidentsial'nosti (na materiale tazovskogo dialekta
sel'kupskogo yazyka) // Acta Linguistica Petropolitana. Trudy instituta lingvisticheskikh
pokazateley (na primere evidentsial'noy sistemy tazovskogo sel'kupskogo) // Voprosy

Other papers:

STRUCTURE OF THE WORK
The dissertation consists of an Introduction, three chapters, a Conclusion, and a list of references. The chapters are devoted to the following issues: 1) Ob-Ugric—Samoyedic contacts; 2) areal and genetic relationships within the Samoyedic language group; 3) narrative strategies in the contact zone formed by the Nganasan language, the Northern and Tym Selkup dialects and the Eastern subdialects of the Vakh Khanty.

MAIN CONTENT OF THE WORK
The Introduction substantiates the relevance of the dissertation research, defines its object and subject, identifies the goals and objectives of the work, characterizes the research methods and the material used, scientific novelty, theoretical and practical significance.

Chapter 1 of the dissertation is devoted to the reconstruction of contacts between the Ob-Ugric and Samoyedic languages. The issues of contacts between the Samoyedic and Yeniseic languages and contacts of the Mansi and Khanty with the Turkic languages are also considered. This supplemental research made it possible to reinforce the conclusions made about the earliest mutual localization of the Ob-Ugric and Samoyedic languages.

1.1. Contacts between Ugric/Ob-Ugric proto-languages and Proto-Samoyedic
The question of the relationships between the (Ob-)Ugric and Samoyedic languages in a historical perspective is considered primarily in the book by E. A. Helimsky “The most ancient
Hungarian-Samoyedic language parallels” (= [Helimsky 1982]). This study can be supplemented by the following material.

1.1.1. Parallels, connecting Ugric and Samoyedic

Based on [Rédei 1988] (if possible, with an updated phonetic reconstruction for Samoyedic and Ob-Ugric), is given an additional list of those lexical parallels, that can be regarded as exclusive Ugric-Samoyedic parallels. There are only four such words, and for none of them it is possible to reliably show that they are common Ugric-Samoyedic innovations, and not an independent preservation of the common Ural lexica.

1.1.2. Ob-Ugric-Samoyed parallels

Ob-Ugric-Samoyedic parallels are divided into a) content words, which can be considered as phonetically exactly corresponding to each other — 12 items; b) content words, which should be interpreted rather as “wandering” terms (phonetic correspondences are irregular even within the languages of the same branch — 11 items; c) terms of spatial orientation — 8 items. Most of the comparisons are also based on [Rédei 1988], but several new etymologies have also been proposed.

1.2. Контакты между мансийским и южносамодийскими (камасинским и селькупским)

In this section, separate correspondences of two types are considered. The most obvious case is represented by such stems, which within the Ob-Ugric branch are presented only in Mansi. More complex are the cases when the word is presented in both Khanty and Mansi, but phonetically the Southern Samoyedic form is not derivable from the Proto-Ob-Ugric or Khanty, while phonetic correspondences with Mansi words look much more natural. As having parallels in Mansi are considered following Selkup words:

Sk. kuə́rpə ‘spruce forest’;
Sk. tapak ‘squirrel’;
Sk. pačə́l’ ‘yellow’ (cf. Proto-Mansi *pəčə́γ ‘deer calf’; the connection between the word for ‘yellow’ and the word for ‘deer calf’ is found also in the Samoyedic languages, cf. TNen tasyexey ‘yellow’ and TNen. təs°ko ‘newborn deer calf’);
Sk. pačiə́ja ‘a spirit living in the hills that howls at night and frightens children’;
Sk. kečə́ / kočə́ ‘servant’ / kečə́ ‘husband’s or wife’s younger brother’;
Sk. piči ‘axe’,
Sk. šä́k ‘salt’ (whatever was the way this originally Iranian cultural term was spreaded in Siberia, it is noteworthy that this exclusive parallel strongly indicates that the Selkup and Mansi
in a certain period were in the same cultural area, which, on the one hand, did not include the Khanty, and on the other hand, did not include Samoyedic languages other than Selkup;

Sk. sırko ‘plant’ (cf. also Kamas šurgu ‘kind of plant, Rus. Dial. Sib. мүхер’ (the roots used to be used to make blue dye for fabric’));

Sk. küm ‘man’ (borrowed from Mans. şum with the same meaning — G. N. Prokofiev also discussed this fact in [Prokofiev 1935: 12]),

sk. konä ‘from the water, from the fire, from the shore to the forest’;

sk. karä ‘to the water, to the fire, from the forest to the shore’;

sk. kolyk/ŋ ‘rarely’;

sk. ʃən ‘almost, barely’.

Several possible Mansi-Kamas parallels are also considered.

1.3. Mansi-Selkup and Mansi-Kamas ethnonymic parallels

Several parallels in ethnonymy can be noted:

Sk. šőső-ʃum / šūššə-ʃum ‘human being’ (the word is of Proto-Selkup origin, it is also used as the ethnonym by speakers of the Ob dialect, as well as within the southern dialect group) has a parallel in Mansi: *šošəy ~ *šošəy > N sossă ~ sâssă [sossa], LM šâsšă ~ šâši ~ šoši, LU šâsi, P šâši, K šâšə ‘local, native’, cf. in particular the following collocations: N taw tǐt âlnē sossă şum ‘he is a native inhabitant here’; P tau tǐt šaši khum, pēš-vuil täkw mōyāt tǐt ŏli ‘he is a local person, he lives on his land from birth’; K šâšə khom, tēlēm mōtē tēt ŏli ‘he is a local person, his place of birth is here’.

Sk. maš kuł ‘a Selkup clan from the Parabel River’ [Prokofiev 1935: 12], where Sk. maš is a good formal correspondence for Mansi *mannči > N mańśi [mâńśi], LM mańś ~ mońś, LU P mańś, K mańś ~ mańś, T mańśi ~ mänči ‘Mansi’, which may indicate the Mansi origin of this Selkup clan.

One can cautiously assume an indirect borrowing of this ethnonym into Kamas (the original word originated from Mansi, but into Kamas it was borrowed from Selkup). If we assume that the clan name maś was transferred from the Selkups to the Kamas (which could happen if a certain group of Selkups switched to the Kamas language), it could explain the ethnonym kyąmāžə ‘Kamasin’ as ‘a representative of the Mas’ family from the Kan River’.

1.4. The history of linguistic areas on Upper Ob

1.4.1. Contacts between Southern Samoyedic and Yeniseic

Contacts between the Samoyeds and the Yeniseic peoples should be recognized as rather late: Kamas and Mator were in contact with Kot, Kamas possibly also with Assan, Selkup - with
Ket. Thus, we see separate contact situations that are predictable from the linguistic landscape of Western Siberia in the 17th century.

The most interesting question concerns Ket-Selkup parallels. Most of them are of late origin, their appearance dates from the period after the dialectal divergence of the Selkup language: as the dictionary [Alatalo 2004] shows, borrowings from Ket are limited predominantly to the Northern and Tym dialects of Selkup. Particularly interesting, however, are the earliest Selkup-Ket parallels (dating from the period before the dialect differentiation of Selkup). These parallels include isomorphic models of compound numerals (see [Kryukova 2013] and also [Urmanchieva 2019a], where it is shown that the Selkup-Ket model is not repeated completely in none of the languages of the region). Within these parallels we also find a significant (compared to the number of content “cultural” words) percentage of “basic borrowings” — the term is used in accordance with [Haspelmath 2008: 5-6, 2009: 46-49] and [Myers-Scotton 2002: 239]. These are such borrowings from Ket into Selkup as basic significant words (‘empty’; ‘straight, correct’; ‘back, back, upside down’), auxiliary words (a particle which is used to form the superlative degree of attributes; a particle which is used to form a comparative / weakened degree of adverbials) and a number of discursive markers (‘enough, stop it’; ‘suddenly’). M. Hasepalmath explains the presence of “basic borrowings” by the prestige of the donor language. Thus, taking into account the character nature of the Proto-Selkup-Ket parallels, we can reconstruct such a contact situation in which the ancestor of Ket was a much more prestigious idiom than Proto-Selkup.

1.4.2. 1.4. The history of linguistic areas on Upper Ob: External Contacts of the Samoyedic and Ob-Ugric Languages.

Speaking of Upper Ob, I refer to the territory of the nowadays “Ostyak” ethno-cultural area (which includes the Khanty, Selkups and Kets) and of the corresponding nowadays linguistic area, formed by the Khanty, Ket and Selkup languages. This area in its modern form, apparently, was formed rather late, since only separate dialects of the two Uralic languages are involved in it: according to the Selkup dictionary [Alatalo 2004], all borrowings from Khanty (as well as from Ket ) are concentrated in the Tym and/or Northern Selkup dialects. From the Khanty side, this contact was also limited to Eastern dialects. Thus, we can regard this contact situation and this linguistic area as relatively young ones.

As it was already mentioned, there was an earlier contact situation, in which the Ket language and Selkup took part before the differentiation of Selkup dialects began. In contrast to borrowings from Eastern Khanty, limited to two Selkup dialects, the Mansi-Selkup parallels discussed above are present in all dialects of the Selkup. These two facts indicate that originally present in the Upper Ob area were the Mansi (in the west), the Selkup (in the center), and the Ket (in the east), but not the Khanty (which perhaps was represented on the periphery of the area). It
is also possible to include in the Upper Ob area the Kamas language — this suggestion can be supported not only by some linguistic facts analyzed in Section 1.2., but also by the results of toponymical research by E. G. Becker. In her work “Selkup toponyms of Western Siberia”, she singled out a group of toponyms of Southern Samoyedic origin on the territory of the modern settlement of the Selkups. Her research allows localizing Kamas in the Chizhapka and Parabel basins and further in the Tom basin. Some of the Turkisms represented in the Kamas language could also appear as a result of the contact of the Kamasins with the Turks in these territories. Thus, the word for ‘salt’ in Kamas has two phonetic variants: tʻoš, tʻuš [Donner, Joki 1944: 73b]. The variant tʻoš could not have been borrowed from the Turkic languages spoken in the Sayan region, but could have been borrowed from Turkic languages such as the language of the Tomsk Tatars.

It is also necessary to briefly dwell on some moments in the history of the Khanty language, which allow us to complete the picture.

If now the border between the Mansi and the Khanty runs from north to south, it seems that formerly this border run from west to east and the Mansi occupied the southern territories, while the Khanty — the northern. The history of Turkisms in the Ob-Ugric languages can serve as evidence in favor of this. An article by A. Kannisto [Kannisto 1925] is devoted to this topic. In this article, A. Kannisto provides a list of Mansi borrowings from the Turkic languages and, in the presence of parallel borrowings into Khanty, also Khanty data. The analysis of these loans allows us to see the following. First, the most obvious fact is that the main influx of borrowings is late borrowings from the Turkic languages into those dialects of Mansi and Khanty that are located on the southern periphery of the (modern) Ob-Ugric world, on the border with the Turkic languages. But outside these two southern groups, the representation of Turkisms in the Khanty and Mansi dialect continuum differs significantly. There are significantly more borrowings in Mansi than in Khanty, which have a wide dialectal distribution and have reached the farthest from the Turkic northern dialect (54 versus 18). In Mansi, in addition, there are significantly more borrowings that did not reach the northern dialect, but at the same time crossed the borders of dialect groups (193 versus 5). Both of these facts indicate on the different periodization of the contacts of the two Ob-Ugric languages with the Turkic languages: in Mansi, a significantly larger number of Turkisms spread in the period preceding the (final) dialect differentiation. Thus, the history of Turkisms in the Ob-Ugric languages, in my opinion, definitely supports the hypothesis that initially the border between the Mansi and Khanty languages run along the west—east line and the Mansi occupied more southern territories (on the border with the Turkic languages), while the Khanty language came into contact with the Turkic languages significantly later.
Thus, I suppose that the formerly Upper Ob linguistic area included Mansi, Selkup and Kamas, on its eastern periphery, the Selkups were in contact with the Kets. The Khanty language was distributed north of the Mansi language and probably occupied the northern periphery of this area. Closest to the Mansi was located the ancestor of the modern Eastern Khanty. The Mansi coexisted with the Southern Samoyeds around in the Parabel / Chizhapka area (where E. G. Becker identified the zone of the Kamas toponymy). During the period of these contacts, in my opinion, Kamas and Selkup were already two separate languages, but the dialect differentiation of Selkup had not yet begun, since those phenomena that I consider as separate Mansi-Selkup areal isoglosses are equally characteristic of all Selkup dialects. Subsequently, the relative position of the Khanty and Mansi languages on the linguistic map of Western Siberia changed: the Eastern Khanty penetrate the Upper Ob area, then the settlement of the Northern and Southern Khanty in a wide territory stretching from north to south along the middle Ob and Irtysh, interrupts the Mansi connection with the Eastern Khanty and finally cuts off the Mansi from the Upper Ob area. During this period, the modern “Ostyak” area began to form in the Upper Ob area, including the Selkup dialects, Eastern Khanty dialects and Ket language.

Chapter 2. Areal and genetic relations within the Samoyedic language group

2.1. Lexicostatistical data for Samoyedic languages

An important topic for the reconstruction of the linguistic prehistory of Western Siberia is contacts within the Samoyedic linguistic community: the reconstruction of these linguistic contacts allows us to build certain hypotheses about the mutual geographical localization of the Samoyedic languages in different historical periods and about the migration routes of the Samoyedic peoples. Obviously, this topic is directly related to the problem of internal genetic classification of Samoyedic languages: without understanding exactly the genetic relationships between languages, we cannot confidently separate secondary, contact induced isoglosses both from the accidental preservation of common archaisms and from common genetic innovations. This chapter therefore begins with the problem of the internal classification of the Samoyedic languages. The primary classification was built by using lexicostatistical method. Within the framework of this study, not the glottochronological aspect of applying this method is important (that is, not an attempt to obtain absolute dating of the separation of languages), but its application to build a classification scheme, (that is, an attempt to measure the relative distances between languages). For lexicostatistical purposes two lists were collected: the 100-word list by M. Swadesh and the 100-word list obtained in the Loanwords Typology project — the so-called Leipzig-Jakarta list, published in [Tadmor 2009: 69–71] (hereinafter the Leipzig list).
Below are the results of the lexical comparison for these two lists. The column with Mator data is visually separated from the rest, since a different system of counting is used for Mator: in particular, if there are synonyms for expressing the same meaning, the presence of a cognate of the Mator word in another Samoyedic language is counted as a match, but the absence of a cognate of one of the synonyms in Mator is not counted as a mismatch. With a rather scarce fixation of the Mator lexics, we cannot reliably interpret the fact of the lack of correspondence: whether it reflects the real absence of a cognate word in Mator or just a gap in our data. The difference in calculations methodic in this case gives higher percentages of correspondences in the Mator language column. Thus, this data can only be compared with other data within that column, but not with data in other columns.

Table 1. Percentage of common vocabulary according to the list of M. Swadesh for Samoyedic languages

<table>
<thead>
<tr>
<th></th>
<th>TNen</th>
<th>FNen</th>
<th>TEn</th>
<th>FEn</th>
<th>Ng</th>
<th>Sk</th>
<th>Km</th>
<th>Mt</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNen</td>
<td>85.08%</td>
<td>69.23%</td>
<td>80.17%</td>
<td>52.13%</td>
<td>47.96%</td>
<td>53.09%</td>
<td>74.19%</td>
<td></td>
</tr>
<tr>
<td>FNen</td>
<td>64.95%</td>
<td>70.68%</td>
<td>50%</td>
<td>49.19%</td>
<td>53.98%</td>
<td>72.04%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEn</td>
<td>67.58%</td>
<td>59.82%</td>
<td>47.96%</td>
<td>54.38%</td>
<td></td>
<td>69.89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEn</td>
<td>56.03%</td>
<td></td>
<td>46.72%</td>
<td>56.75%</td>
<td></td>
<td>70.65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ng</td>
<td>40.8%</td>
<td></td>
<td>35.57%</td>
<td></td>
<td></td>
<td>56.98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63.15%</td>
<td></td>
<td>63.44%</td>
<td></td>
</tr>
<tr>
<td>Km</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70.63%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Percentage of common vocabulary according to the Leipzig list for Samoyedic languages

<table>
<thead>
<tr>
<th></th>
<th>TNen</th>
<th>FNen</th>
<th>TEn</th>
<th>FEn</th>
<th>Ng</th>
<th>Sk</th>
<th>Km</th>
<th>Mt</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNen</td>
<td>82.92%</td>
<td>62.80%</td>
<td>75.40%</td>
<td>50.41%</td>
<td>41.98%</td>
<td>51.30%</td>
<td>73.33%</td>
<td></td>
</tr>
<tr>
<td>FNen</td>
<td>64.46%</td>
<td>70.73%</td>
<td>49.16%</td>
<td>44.69%</td>
<td>51.69%</td>
<td>74.44%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEn</td>
<td>81.35%</td>
<td>62.18%</td>
<td>45.73%</td>
<td>53.04%</td>
<td></td>
<td>70.78%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEn</td>
<td>54.16%</td>
<td>43.41%</td>
<td>52.58%</td>
<td></td>
<td></td>
<td>73.56%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ng</td>
<td>40.625%</td>
<td></td>
<td>52.21%</td>
<td></td>
<td></td>
<td>60.67%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sk</td>
<td></td>
<td></td>
<td>60%</td>
<td></td>
<td></td>
<td>62.22%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Km</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>71.59%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First of all, we see that both lists, despite the difference in specific numerical values, give the same results in terms of relative distances between pairs of languages. The most closely related languages according both lists are: Tundra Nenets and Forest Nenets and Tundra Enets and Forest Enets: only within each of these pairs of languages is the percentage of vocabulary matching higher than eighty. The identification of these two pairs, of course, is not so much based on the lexicostatistical data, but rather confirms their correctness: the fact that within the Samoyedic branch the Tundra Nenets is closest to the Forest Nenets, and the Tundra Enets is closest to the Forest Enets, is a completely trivial knowledge. According to the percentage of correspondences, the Nenets-Enets subgroup is also easily identified.

As shown by E. A. Helimsky on lexicostatistical calculations for the Finno-Ugric languages, analyzed in detail in [Helimsky 1982: 11–27], Finno-Ugric languages that belong to different subgroups but came into contact with each other, give an higher percentage of lexical
correspondences even within the Swadesh list, which is explained by their specific “areal-genetic” relations. The same “areal-genetic” relations, that cross the subgroups boundaries can also be identified on the basis of lexicostatistical data within the Samoyedic languages. It can be seen that within the “classical” set of Northern Samoyedic languages, the percentage of lexical correspondences was significantly influenced by contacts. Let's start with the two Nenets languages. Since both Nenets form a low-level subgroup, the distance from them to any of the Enets languages (forming another low-level subgroup) should be approximately the same. However, the percentage of coinciding vocabulary between Tundra Nenets and Forest Enets is higher than between Tundra Nenets and Tundra Enets. This mismatch, of course, also reflects a rather well-known fact: the contacts between the Tundra Nenets and the Forest Enets. The percentage of coinciding vocabulary between Forest Nenets and Forest Enets is also higher than between Forest Nenets and Tundra Enets. But the discrepancy in this case is smaller. And the higher percentage of vocabulary coincidence between Forest Enets and Forest Nenets in this case does not necessarily indicate direct contact between these languages. If we take into account a high percentage of vocabulary coincidences between Forest and Tundra Nenets, it becomes obvious, that almost any contact induced change in Forest Enets lexics under the influence of Tundra Nenets automatically makes the Forest Enets more similar to the Forest Nenets (without direct contact with the latter).

The percentage of common vocabulary for Tundra Enets and Nganasan (62.18% according to the Leipzig list) is significantly higher the percentage of matches between Forest Enets and Nganasan (54.16% according to the Leipzig list). In this case, too, the higher percentage of correspondences reflects the situation of linguistic contacts: a well-known fact is also the contacts between these two languages.

Nganasan demonstrates an increased percentage of correspondences not only with the Tundra Enets, but also with the Forest Enets. It was discussed above that in the case of an increase in the percentage of correspondences between Forest Nenets and Forest Enets, we do not draw a conclusion about direct contacts between languages: in that case, Forest Enets, changing under the influence of Tundra Nenets, automatically became more similar to the closest relative of Tundra Nenets — to Forest Nenets. However, we cannot explain the increase in the percentage of correspondences between Nganasan and Forest Enets as a «mirroring» of contacts between Nganasan and Tundra Enets. The fact is that in the case of contacts between Tundra Nenets and Forest Enets, the direction of influence was from Nenets (as the donor language) to Enets (as the recipient language). But in the case of contacts between Tundra Enets and Nganasan, there is no reason to consider Tundra Enets to be the donor language. But only such a situation (Tundra Enets — donor, Nganasan — recipient) would explain the automatic, “mirrored” increase in the
percentage of correspondences between Nganasan and Forest Enets in the absence of direct contacts. Thus, it is more logical to assume that before the divergence of Proto-Enets into Forest and Tundra Enets, there was contact between (Proto-)Nganasan and Proto-Enets. And after the divergence of Proto-Enets, contact continued only with the Tundra Enets.

As for Kamas and Selkup, the percentage of matches between them (63.15% for the Swadesh list, 60% for the Leipzig list) is higher than the percentage of matches between any of these two languages and any Northern Samoyedic. On the one hand, the highest percentage of matches between Kamas and Selkup (compared to the distance from any of these languages to any of the Northern Samoyedic languages) suggests that these two languages do constitute a separate Southern Samoyedic branch. On the other hand, in this case, we would expect that the distance from both Selkup and Kamas to any of the languages of the other branch would be approximately the same, while for Kamas the percentage of common vocabulary always turns out to be higher than for Selkup. I assume that Selkup and Kamas do belong to the same branch, and the reduced percentage of common vocabulary in the case of Selkup may be a fact of the individual history of this language: Selkup may give lower figures due to the high number of words (maybe, of substrate origin) that do not have reliable Samoyedic parallels (see the next section).

According to both lists, Mator gives the highest percentage of matches with both Nenets languages, than — with both Enets languages and with Kamas (with a minimal difference). I believe that the resulting picture should be interpreted as follows: in the absence of obvious long-term contacts between Mator and Northern Samoyedic languages, and in the presence of obvious opportunities for contacts between Mator and Kamas (both of these facts follow at least from the modern geographical location of the languages), the percentage of correspondence between Mator, Nenets and Enets should be considered as reliably reflecting genetic relationships, and the percentage of matches with Kamas should be considered as increased due to secondary contacts. This preliminary confirms the conclusion of E. A. Helimsky that Mator belongs to the group of Northern Samoyedic languages, and within the Northern Samoyedic languages it is (also in accordance with the assumption of E. A. Helimsky) closer to Nenets and Enets than to Nganasan. This hypothesis is also supported by the analysis of lexical parallels between Mator and different Samoyedic languages, analyzed in 2.3.1.

2.2. Substitution of lexics in two 100-word lists and languages with a reduced percentage of matches with other languages

Both lists show that Nganasan has a relatively low percentage of lexical matches with all Samoyedic languages. It was also already noted above that while Kamas and Selkup constitute a separate South Samoyedic branch, it turns out that the percentage of lexical correspondences with Samoyedic languages outside this branch is lower for Selkup than for Kamas. The low percentage
of matches can have two main explanations: either languages with a low percentage of matches were separated from the protolanguage earlier than all the others, or for some reason the replacement of vocabulary in these languages occurred more intensively. In case of Samoyedic languages the first explanation does not seem reasonable to me because. Firstly, there seem to be no reason to exclude Nganasan from Northern Samoyedic branch (see [Gusev 2016] for a convincing argumentation) Secondly, the percentage of lexical correspondences between Kamas and Selkup also quite convincingly testifies to their unification into one branch: in this case, there are also no sufficient grounds to postulate a specifically early separation of the Selkup from the Proto-Samoyedic. As for the second explanation, for Nganasan it is a priori confirmed by the common and quite fair (although, it seems, nowhere factually substantiated) opinion that Nganasan has an abnormally high percentage of non-etymologized vocabulary.

In this work, both 100-word lists are marked as follows: for each language, unique (not uniting it with any other Samoyedic language, that is, reducing the overall percentage of lexical correspondences) lexical units of two types are singled out: a) unique items, which have no Samoyedic etymology but at the same time not being borrowed from any of the known languages; b) unique items, which do have a Samoyedic etymology (but the lexical substitution in question occurring only in this language).

Table 3 summarizes the data on unique lexical items of both types in the Swadesh and Leipzig lists. The second column of the table shows the number of unique internal substitutions, and the third column shows the number of non-etymologizable words. The fourth column gives the total number of unique lexical items.

Table 3. Unique lexical items in Samoyedic languages (in the Swadesh and Leipzig lists)

<table>
<thead>
<tr>
<th>Language</th>
<th>Unique internal substitutions</th>
<th>Unique non-etymologizable words</th>
<th>Total number of unique lexical items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Nenets</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Common Nenets + Forest Nenets</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Common Enets</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Common Enets + Tundra Enets</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Nganasan</td>
<td>11</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>Selkup</td>
<td>9</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Kamas</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Mator</td>
<td>7</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>
The data in the third column of the table are especially indicative: in Selkup and Nganasan, the percentage of non-etymologized vocabulary is significantly higher than in other Samoyedic languages — even within the two lists claimed to represent the most stable part of vocabulary. The numbers in the fourth column also show that exactly the same two languages demonstrate the significantly higher number of words that have no correspondence (either material or functional) in any other Samoyedic language.

Firstly, the data in Table 3 explain quite well the low percentage of lexical matches that Nganasan has with other Samoyedic languages. Secondly, these data make it possible to substantiate the hypothesis formulated earlier: Kamas and Selkup do constitute a single Southern Samoyedic branch, and this is not contradicted by the fact that the percentage of correspondences between Kamas and other Samoyedic languages is always higher than the percentage of correspondences between the Selkup and other Samoyedic languages.

However, the data in Table 3 allow to formulate yet another question. Selkup and Kamas as the languages of the Southern Samoyedic branch, in terms of distance from each of them to any of the Northern Samoyedic languages, theoretically approach the “ideal” picture of the percentage of lexical correspondences between the languages of different subgroups in case of the absence of contacts: Selkup, like Kamas, turn out to be approximately equidistant from any of the languages from another subgroup, that is, from any Northern Samoyedic language. However, this should not have happened in the case of Nganasan: as the calculations just carried out have shown, Nganasan has a high percentage of words that do not have correspondences in other languages. Accordingly, Nganasan should have a reduced percentage of matches with Kamas and Selkup. Thus, the following paradox is observed: knowing that Nganasan has a high percentage of words that do not have correspondences in other languages, we have to look for a special explanation for the fact that Nganasan (in comparison with other Northern Samoyedic) DOES NOT HAVE a reduced percentage of lexical matches with Selkup and Kamas. In other words, the approximately EQUAL distance from Kamas and Selkup to any of the Northern Samoyedic languages actually indicates an INCREASED percentage of lexical correspondences between Nganasan and Selkup and Nganasan and Kamas. Thus, we observe not a (pseudo) ideal picture of the absence of contacts between the languages of different subgroups, but just evidence of a possible contact between Nganasan and Selkup and Nganasan and Kamas. This question is investigated in details in Section 2.3.

2.3. Additional investigation of the “problem nodes” of the classification using the method of exclusive parallels to determine the common genetic or areal innovations

In this section, lexical parallels are considered, confirming the preliminary conclusions that were made on the material of two lexical lists: the Swadesh list and the Leipzig list. Since these
lists largely overlap, these preliminary conclusions are drawn on relatively modest material. This section presents the results obtained by the examination of all lexical material available for several chosen languages.

The following lists of exclusive lexical parallels have been compiled:

1. Correspondences between Mator and Selkup, Mator and Kamas and Mator, Selkup and Kamas, as well as correspondences between Mator and Northern Samoyedic languages in all possible combinations. This study allows us to answer the question of whether Mator should be considered as belonging to the Northern or to the Southern subgroup. The difference between this study and [Helimski 1997 Ms. (2022)] lies in the fact that the identified parallels are characterized in terms of whether they are archaisms or innovations (and if they are innovations, then whether these innovations are of a genetic or of contact nature).

2. List of Nganasan-Selkup, Nganasan-Kamas and Nganasan-Kamas-Selkup lexical parallels. Nenets-Selkup, Enets-Selkup, Nenets-Kamas and Enets-Kamas lexical parallels are also collected to show, that their number is incomparably smaller.

3. List of Nganasan-Enets lexical parallels.

4. List of Nganasan—Tundra Enets lexical parallels.

2.3.1. Lexical parallels of Mator with Northern Samoyedic languages and Mator with Southern Samoyedic languages as a mean of determining the place of Mator within the genetic classification of Samoyedic Languages

Table 4 shows a quantitative analysis of the identified separate parallels between Mator and Southern Samoyedic languages and Mator and the Northern Samoyedic languages. Counted were a) those units that can be more or less reliably defined as general genetic innovations and b) total number of genetic innovations plus with units that are not reliably defined either as archaisms or as innovations (which, in my opinion, allows us to evaluate general “density” of lexical connections between languages). In other words, parallels that have been identified as secondary areal innovations, as well as units that are undoubtedly identified as common archaisms (uralisms), are excluded from the calculations in any case.
Table 4. Lexical parallels between Mator and Southern Samoyedic languages and Mator and the Northern Samoyedic languages

<table>
<thead>
<tr>
<th></th>
<th>Number of items, that can be regarded as common genetic innovations, that Mator has n common with …</th>
<th>Number of items, that can be regarded as common genetic innovations and parallels, than can not reliably be identified as archaisms or innovations, that Mator has n common with …</th>
</tr>
</thead>
<tbody>
<tr>
<td>… with Kamas and Selkup</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>… with Selkup</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>… with Kamas</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>… with Southern Samoyedic (total number)</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>… with Nenets, Enets and Nganasan</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td>… with Nenets and Enets</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>… with Enets</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>… with Nenets</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>… with Northern Samoyedic (total number)</td>
<td><strong>34</strong></td>
<td><strong>64</strong></td>
</tr>
</tbody>
</table>

As can be seen from this table, the number of parallels that can be interpreted as common genetic innovations is significantly (three times) higher between Mator and Northern Samoyedic languages than between Mator and Southern Samoyedic languages; the general “density” of lexical connections is also significantly (three times) higher between Mator and Northern Samoyedic languages than between Mator and Southern Samoyedic languages. Thus, I (following E. A. Helimsky) define Mator as a Northern Samoyedic language. E. A. Helimsky notes the greater proximity of Mator to the Enets-Nenets subgroup, this is also confirmed by my data. It is interesting, however, to note that further on Mator continues to maintain a distinct proximity to Nenets.
2.3.2. Nganasan-Selkup-Kamas parallels

The number of separate lexical parallels connecting Enets with Southern Samoyedic (no parallels were found), both Nenets languages (two parallels were found) and Forest Nenets separately (two parallels were also found) was compared with the number of separate lexical parallels connecting Nganasan with both South Samoyedic languages (15 parallels are identified, and most of them can be interpreted as innovations of a secondary, areal nature). Obviously, the number of Nganasan-Selkup-Kamas parallels is statistically significant and may indicate special areal relations between these three languages.

2.3.3. Nganasan-Kamas parallels

In this work 25 Nganasan-Kamas lexical parallels are found, 14 of them can be identified as innovations; none of the remaining is reliably identified as archaism. This number is obviously significant if we compare it with the number of separate parallels linking Kamas with Common Nenets (3), with Common Enets (1) and with Forest Nenets (2). Both Forest Nenets—Kamas lexical parallels denote climatically limited artifacts or natural phenomena, so that they can equally be both late terms of contact origin and archaisms, which were not preserved in other Samoyedic languages, whose speakers moved to a different climate zone.

2.3.4. Nganasan-Selkup parallels

In this work 49 Nganasan-Selkup separate parallels were found, of which:

- 3 are defined as archaisms (two phenomena from the field of historical phonetics and one lexical parallel);
- 12 cannot be defined as archaism or innovation;
- 34 are defined as innovations.

Other Samoyedic languages demonstrate a different picture:

- 12 Selkup-Nenets separate parallels, of which:
  - 1 word is undoubtfully defined as archaic;
  - 2 words denote animals, whose names with a high probability could be lost in Enets and Nganasan during the migration of their speakers to the northern territories — I attribute these words to the group of archaisms;
  - 3 words could have a wider distribution: it has been demonstrated that the corresponding lexemes could be replaced in Enets and Nganasan by words of later origin, forming Nganasan-Enets contact isoglosses — I attribute these 3 words to the group of archaisms;
  - 2 words can be defined as innovation;
  - 4 words cannot be defined as archaisms or innovations.

- 4 Selkup-Enets separate parallels, of which:
- 2 words are defined as innovation;
- 2 words cannot be defined as archaisms or innovations;
- 12 Selkup — Forest Nenets separate parallels, of which:
  - 4 words cannot be defined as archaisms or innovations;
  - 8 words with a high probability are identified as common cultural terms that represent secondary contact isoglosses. However, it is also possible that at least some of these terms were originally more widespread, but were later lost in those Samoyedic languages that left this climatic (and, accordingly, cultural) area.
2.3.5. Parallels between Nganasan and Southern Samoyedic languages: a summary of the data
2.3.5.1. Exclusive Nenets-Enets-Selkup-Kamas, Nenets-Enets-Kamas and Nenets-Enets-Selkup parallels
Theoretically, it can be assumed that Enets and Nenets have few separate parallels with Selkup and Kamas (in comparison with the number of Nganasan-Selkup and Nganasan-Kamas parallels) because these languages split later from each other than Nganasan from Proto-Northern Samoyedic and, accordingly, Enets and Nenets did not have time to “accumulate” a difference in vocabulary during the period of separate existence, comparable to the difference between any of them and Nganasan. Thus, the comparable period for Nenets should include both the period of existence of the Nenets-Enets language and the period of the separate existence of Nenets. Therefore, in this section, the Nenets-Enets=Selkup-Kamas, Nenets-Enets=Kamas and Nenets-Enets=Selkup parallels were also briefly characterized, and their number was compared with the number of Nganasan-Selkup-Kamas, Nganasan-Kamas and Nganasan-Selkup parallels, respectively, see Table 5.

Table 5. Separate parallels between Kamas, Selkup and Northern Samoyedic (without archaisms)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Km.-Sk.</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Km.</td>
<td>25</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Sk.</td>
<td>46</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>11 (21)</td>
</tr>
</tbody>
</table>

It is easy to see that the number of Selkup-Kamas-Nganasan parallels of this kind is three times higher than the number of parallels with Kamas and Selkup that could “accumulate” over the corresponding period of its history the Nenets language (16 versus 5), in the case of parallels with Kamas, for Nganasan this number is five times higher than for the Nenets (25 versus 5), in the case of parallels with Selkup, even with the most “favorable” calculation method for the
Nenets, the number of such parallels for the Nganasan is more than twice as high as for the
Nenets (46 versus 21).

Table 6 takes into account only those words for which it is possible to show their innovative
character. Here the difference is even more significant: the number of Selkup-Kamas-Nganasan
innovations is 12 times higher than the number of parallels with Kamas and Selkup that could
“accumulate” over the corresponding period of its history the Nenets language (12 versus 1), in
the case of parallels with Kamas, for Nganasan this number is seven times higher than for the
Nenets (14 versus 2), in the case of parallels with Selkup, the number of innovations for the
Nganasan is eight times higher than for the Nenets (34 versus 4).

Table 6. Separate parallels between Kamas, Selkup and Northern Samoyedic (innovations only)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Km.-Sk.</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Km.</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sk.</td>
<td>34</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Thus, the number of separate lexical parallels between Nganasan and Southern Samoyedic
is significantly higher than the number of the same parallels for Nenets and Enets. More important,
it has been shown that most of the parallels between Nganasan and Southern Samoyedic can only
be interpreted as secondary contact innovations. Thus, the fact of historical contacts between the
Nganasan and South Samoyedic languages is proven.

2.3.6. Nganasan-common Enets parallels

As already mentioned, lexicostatistical data show that genetically Enets is undoubtedly
closest to Nenets and Mator. However, they also show that even within the most stable basic
vocabulary, Nganasan gives an increased percentage of matches with Enets (and also, separately,
with Tundra Enets). While the contacts of the Tundra Enets and Nganasan are well attested
historically, we know nothing about separate contacts between the Proto-Enets and Nganasan. In
this section, as an illustration, an incomplete list of 60 Nganasan-common Enets parallels is given.

2.3.7. Parallels between Nganasan and Tundra Enets

In this section, as an illustration, a list (also incomplete — 25 units) of words that are
recorded only in Nganasan and Tundra Enets is given.

2.4. Areal and Genetic Relationships within the Samoyedic languages: Conclusions

Thus, it is possible to propose a scheme of genetic classification of the Samoyedic languages.
Further, this scheme can be supplemented with information about different-time contact situations. Scheme 2 represents the contacts of Nganasan with Southern Samoyedic languages.

Scheme 3 represents later contacts of Nganasan with Proto-Enets and, at the same time, not too intense contacts of Mator simultaneously with Kamas and Selkup.
Scheme 3. Nganasan—Proto-Enets and Mator-Selkup-Kamas contacts

Scheme 4 represents the latest contacts: Tundra Nenets with Forest Enets, Nganasan with Tundra Enets, Mator with Kamas.

Scheme 4. Historically attested contact situations within the Samoyedic language group

2.5. Reconstruction of several consecutive states of linguistic map of Western Siberia

Taking into account the proposed scheme of the divergence of the Samoyedic languages together with the reconstructed contact situations of several time depth (contacts within the
Samoyedic languages, reflected in the three diagrams above, and the external contacts of the Samoyedic languages with the Ob-Ugric languages discussed in the first chapter of the dissertation, including separate contacts between Mansi and Selkup), it was possible to reconstruct a number of migration processes, starting from the period of the divergence of the Proto-Samoyedic and up to the period of historically attested distribution of Samoyedic languages. I support the view that the Samoyedic ancestral homeland was located “in the region between the middle Ob and Yenisei, approximately around the triangle «Tomsk — Krasnoyarsk — Yeniseisk»” [Helimsky 1989 / 2000: 21].

The reconstruction of the first stages immediately after the first divergence of the Proto-Samoyedic language should reflect the following phenomena:

- the of contacts between the Selkup and, to a lesser extent, the Kamas with the Mansi language;
- early separation of Nganasan from the northern Samoyedic branch and its separate contacts with Kamas and Selkup;
- belonging of Mator to the Northern Samoyedic branch (within one taxonomic unit with Nenets and Enets and further — with Nenets).

The reconstruction of the migration routes is also directly related to the issue of the reconstruction of the material culture of the speakers of Proto-Samoyedic, since a better understanding of this issue will make it possible to better understand the causes of such migrations. E. A. Helimsky states, that one of the possible reasons for the divergence of the Samoyedic proto-language and for northward migration of peoples speaking the ancestor languages of Nenets, Enets and Nganasans, was the increased mobility of the population due to the development of transport reindeer breeding among the Samoyeds. But in the taiga zone there are no material cultures, in which the provision of human life would be carried out exclusively at the expense of reindeer herding; reindeer breeding here can only be of an auxiliary character, of secondary importance (at least due to the lack of a sufficient amount of reindeer moss). Cf. the situation among modern Agan Nenets: herds of domestic deer are small, semi-free grazing suggests the possibility of conducting other activities — hunting and fishing ([Zenko-Nemchnova 2006]). Accordingly, the question arises: what type of culture could be characteristic of Proto-Samoyeds at least at the beginning of migration? In general, it can be assumed that whatever was the change in the population of the taiga zone of Western Siberia in the last couple of millennia, the environment should assume the reproduction of approximately the same economic and cultural complex, which is most consistent with the biodiversity and landscape of this region. If we look at the modern peoples inhabiting the taiga zones between the Ob and Yenisei rivers, we realize that the main component of their culture is fishing. It is combined with hunting, but, unlike the Evenks inhabiting
the right bank of the Yenisei, for whom hunting is the main core of the traditional culture, among the peoples living to the left of the Yenisei, the traditional material culture is focused specifically on fishing (The role of hunting to some extent increases in the northern zone of the taiga, near the forest-tundra zone, but to a large extent it was caused by need to meet the demands of Russian fur trading. And in any case now we are interested in more southern cultures, geographically closer to the reconstructed Samoyed homeland.) It can be summarized as follows: in the southern part of Western Siberia taiga zone there is not a single people who would not have developed fishing as a central component of traditional culture. The statement that fishery was and is a characteristic feature of Western Siberian traditional cultures and the statement that regardless any possible changes of languages and ethnic groups in this territory the same economic and cultural complex was reproduced can also be confirmed by the fact that “researchers attribute the emergence of weirs fishing to the end of IV–III millennium BC; in the Bronze Age it was widely used by the population of the West Siberian taiga (Vasiliev, 1962a, pp. 142–143). Numerous finds of sinkers testify to the traditions of fishing with nets (Kiryushin, Maloletko, 1979, p. 128; Kosarev, 1981, p. 21)” [Golovnev, Perevalova 2005: 79].

For the Samoyedic proto-language, a significant number of terms related to fishing can be reconstructed: *kəŋår ‘fish-trap’, *wikər1 ‘gaff’, *juŋ ‘fishing weir’, see [Helimsky 1989 / 2000: 16], also *poŋkə ‘net’ [SW 127], *putə ‘net string’: Sk. pûta ([Alatalo 2004: 458], where the Nenets correspondence is indicated), TNen pud9, TEn. piðo, Ng. hitə. Thus, for the Proto-Samoyeds, we can reconstruct fishing as an essential component of traditional culture. Small rivers (including the upper reaches of relatively large tributaries of the main water arteries) are most suitable for weirs fishing. And in general, the “fishing cultures” of Western Siberia are characterized not so much by their orientation towards the course of the region’s major river, but rather by the settlement of local groups within the basins of its tributaries, along the set of small rivers (cf. the situation with the Khanty, Mansi and Selkup dialects: each of them was formed in the basin of one or another medium-sized river with its tributaries). Precisely this way of settlement E. A. Helimsky reconstructed for the Selkups after they left the Samoyedic homeland. He stated, that the speakers of the Proto-Selkup initially settled down around the upper reaches of the Ket and Chulym and then, gradually moving downstream in the basins of these rivers, came to the Ob. Assuming that the speakers of the dialects of Proto-Samoyedic immediately after its divergence all were the representatives of the more of less identical riverine culture, I consider it fair to reconstruct the same principle of settlement not only for the ancestors of the modern Selkups, but also for speakers of other dialects of the Samoyedic proto-language.

Thus, I assume that the first waves of migrations from the Samoyedic homeland were as follows. First of all, all migrations were oriented to the Ob basin, and not to the Yenisei basin.
(neither the Mansi-Selkup nor the Nganasan-South Samoyedic contacts can be explained otherwise). The speakers of Proto-Northern Samoyedic moved northwards, approximately to the Tym basin, then the ancestors of the Nganasans remained there, and the speakers of Proto-Nenets-Enets-Mator continued the northward migration till the Vakh basin. The speakers of the Kamas and Selkup after the separation of the North Samoyedic group probably remained for some time in the territory of the homeland, but later split up and dispersed along the tributaries of the Ob: the ancestors of the Kamas probably went along the left-bank tributaries and eventually came to the Parabel and Chuzik (in the basin of these rivers E. G. Becker identifies a number of hydronyms of Kamas origin). As a result of this migration, having reached Parabel and Chuzik, the speakers of the Proto-Kamas reached the zone of contacts with the Mansi language (as was shown in the first section of the dissertation, it can be assumed that initially the division of territories between the Mansi and the Khanty did not take place along the “north-south” line, as now, but along the “west-east” line, whereas the Mansi occupied more southern territories, and the Khanty more northern ones), see Figure 1.

*Figure 1. The first waves of migration from the Samoyedic homeland*
Further, the Selkup ancestors settled down around the upper reaches of the Chulym and Ket migrated down these rivers to the Ob. As a result, they reached the territories inhabited by the ancestors of Kamasins on the left bank of the Ob. Then the Selkups pushed the Kamasins from the left-bank tributaries of the Ob, from the Parabel and Chuzik, and subsequently Selkups enter the zone of contacts with the ancestors of the Mansi.

In the same period, the northern Samoyeds continued to develop the right tributaries of the Ob: the ancestors of the Nganasans remain in the Tym basin: it is their rather long stay in this zone that makes it possible to explain the contacts of the Nganasan with the Kamas and the Nganasan with the Selkup. The ancestors of the Enets remain on the Vakh, the speakers of the Nenets-Mator proto-language develop the territories from the Agan to the Lyamin on the right bank of the Ob (that is, approximately the territories on which the Forest Nenets later remain); speakers of the Proto-Mator then moved to the left bank (vis-a-vis the Nenets territories), see Figure 2.

*Figure 2. Settlement of the Samoyeds in the Ob basin*

This distribution of proto-languages was in a later era disrupted by the expansion of the Khanty eastwards, along the Middle Ob: to the Vasyugan, up the Ob approximately from the mouth
of the Irtysh to the mouth of the Vakh, and later — to the Vakh. The continuing migration of Khanty changed the initial mutual localization of the speakers of Proto-Mansi and Proto-Khanty: settling north along the Ob and south (along the same line) along the Irtysh and in the Vakh and Vasyugan basins in the east, the speakers of Proto-Khanty interrupted contacts between the Mansi and the Selkups. As a result, the Mansi retreated to the west and then to the north, to the basin of the Northern Sosva and beyond.

At the same time, the eastward migration of the Khanty along the Ob and further, to the basins of the Vasyugan and then the Vakh, causes a number of chain movements in the location of the speakers of Samoyedic proto-languages. Firstly, this migration “splitted” the Nenets-Mator dialect continuum, pushing the speakers of the Proto-Nenets dialects from the Ob northeast, even closer to the zone of the modern distribution of the Forest Nenets, and the speakers of the Proto-Mator (who were on the left bank of the Ob), to the more southern territories, probably in the Vasyugan basin. In this regard, it is important to mention the Khanty name of the Nyurolka, the right tributary of the Vasyugan, recorded in the 18th century by G. F. Miller as Iaring-jach-jogon [Ehlert 1996: 196] and in the 19th century by M. A. Castren as Jargan-jogan [Castren 1860: 271]. This toponym means ‘Nenets river; Samoyedic river’. The considered ethnonym is reconstructed for the Proto-Ob-Ugric in the form *jărγən / *jăγrən [Napolskikh 2000/2018: 216] and denotes ‘the Nenets’. Obviously, we cannot assume the presence of the Nenets on the right bank of the Vasyugan; the association of this ethnonym with the Selkups is also excluded, see [Tuchkova 2014: 65]. Since the division of Mator and Nenets is the last division within the Northern Samoyedic linguistic community, it can be assumed that the Mators were close to the Nenets not only in language, but also in culture. In this case, the possibility of applying the same exoethnonym to two groups — Nenets and Mators — is quite likely. Thus, this “strange” Khanty toponym makes it possible to localize the Mators in the basin of the Nyurolka, the right tributary of the Vasyugan, which confirms the reconstructed migration route. From there, the speakers of the Proto-Mator passed along the southwestern periphery of the Samoyedic area and finally become the immediate western neighbors of the Selkups and Kamasins. Further, the Kamasins, together with the Mators, continue to move to the east (perhaps just to the lower reaches of the Tom, where E. G. Bekker notes another area of Kamas toponyms). From here, finally, they are forced out to the Sayans by a late wave of migrations of speakers of Turkic languages, which subsequently swallowed up speakers of southern Khanty dialects, and speakers of southern Mansi dialects, and southern Yeniseic languages and speakers of Sayan-Samoyedic languages.

The migration of the Khanty to Vasyugan also causes the exodus of a part of the Vasyugan Selkups to the Tym; this is how, in my opinion, the Central Selkup dialect area was formed. At the same time, having gone to the Tym, the Selkups displace in the northeast direction the Nganasans,
who till that period continued to inhabit the Tym basin. In the same northeast direction moved also the group of Enets (occupying an intermediate position between the Nenets Nganasans). Proto-Enets—Nganasan contacts may continue during this period. Thus, the northern Samoyeds actually occupy the starting positions, from which the well-known late ethnic processes are naturally explained: the formation of the Forest Nenets in the Agan and Pur basins and the Tundra Nenets in the Arctic tundra zones, the formation of the Forest Enets in the basin of the middle Taz and the Tundra Enets in the lower reaches of the Yenisei, migration of the ancestors of the Nganasans to Taimyr and where the final formation of this ethnic group took place. These migration processes are mapped in Figure 3.

*Figure 3. Changes in the linguistic map caused by the expansion of the Khanty*

Interestingly, it is possible to establish a certain correlation between the contact situations reconstructed on the basis of linguistic material and the types of material cultures identified in [ИЭАС 1961]. Thus, ethnographic investigations provide a fairly definite picture of the relationships between the Nganasan material culture and the material culture represented by the Kets and Selkups (but not by the Enets and Nenets). In accordance with [ИЭАС 1961], the
Nganasan material culture is most like the Selkup material culture in the following areas: types of clothing (section written by N.F. Prytkova, map on p. 325), types of headgear (section written by N.F. Prytkova, map on p. 368), types of boats (section written by V. V. Antropova, map on p. 129). In addition, in ИЭАС 1961 reindeer transport, draft dog breeding (not for Nganasans), skiing (not for Nganasans), dwellings, ornaments, and shaman tambourines are also considered. According to this last parameter, the Nganasans and Enets form a special type; they also have a special “Taimyr” type of clothing, combined with clothing of the Yenisei type by Nganasans; as for dwellings, they naturally use a tent, a dwelling common to many peoples, but the Nganasans have a special, original construction for joining the main tent poles, which allows varying the size of the tent. Thus, according to a number of parameters, the Nganasans are united with the Yenisei peoples (Kets and Selkups), according to other parameters they form a separate type, delimiting them (sometimes together with the Enets) from the Kets and Selkups, but equally from the Nenets. Thus, the ethnographic material confirms the existence in the historical perspective of both Nganasan-Selkup and Nganasan-Enets contact situations.

Chapter 3. Narrative strategies in the contact zone formed by the Nganasan language, the Northern and Tym Selkup dialects and the Eastern subdialects of the Vakh Khanty

This chapter is devoted to a later area that existed in the period immediately preceding those migrations that formed the linguistic map of the north of Western Siberia in its modern shape. This zone included Nganasan language, the Northern and Tym Selkup dialects and the Eastern subdialects of the Vakh Khanty. It should be noted, that (in consistence with the map I reconstructed) there was almost certainly no direct contact between the Nganasan language and the Vakh Khanty dialects (we do not have any linguistic or historical evidence of Nganasan-Khanty contacts). So we should assume a chain transmission of common phenomena in this area: Selkup, which was in the center of the chain, simultaneously contacted both with Khanty dialects and with Nganasan. In this contact zone, a typologically rare system of discursive strategies was formed, which are used in folklore and personal narratives.

The discursive functions of verbal forms in narrative are traditionally understood as such use of aspectual, temporal or evidential forms, which is determined not by their grammatical semantics, but by the need to mark a certain position of one or another text fragment in the overall structure of the narrative (cf. the works in which this phenomenon is analyzed on material of various languages, in [Gusev et al. (ed.) 2008]). We can distinguish between the following main discursive functions of verb forms:

- sequentia function: the situations constituting the main story line are described in chronological order one after another;
- introductive function: as a rule, special verbal forms are used in the initial, introductory fragment, where the main characters are introduced, their “personal data” is introduced, a brief background of the main plot can also be stated.

However, also another interesting dimension of the use of verb forms in discourse can be touched upon. Thus, verb forms in narrative may have another function, namely, the use of different verb forms can distinguish between texts of different genres. Different types of discourse can also be distinguished in the so-called narrative mode; it is the opposition of texts of various genres, such as personal narratives (lifestories) or folklore narratives (with possible further subdivision).

In several idioms (the Tym Selkup, the Middle Taz variety of the Northern Selkup, the Eastern varieties of Vakh Khanty and Nganasan), a typologically unique system of marking the ternary genre opposition (fairy tales — historical legends — personal narratives) was formed.

Table 7. Idioms with ternary opposition of narrative strategies

<table>
<thead>
<tr>
<th></th>
<th>Tym Selkup</th>
<th>Northern Selkup</th>
<th>Nganasan</th>
<th>Vakh Khanty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical legends</td>
<td>Introductive</td>
<td>Inferentive, Narrative Reportative</td>
<td>Narrative Reportative</td>
<td>Reportative</td>
</tr>
<tr>
<td>Sequential</td>
<td>Habitual</td>
<td>Narrative Reportative</td>
<td>Reportative</td>
<td>-∅</td>
</tr>
<tr>
<td>Fairy tales</td>
<td>Introductive</td>
<td>Habitual</td>
<td>Inferentive, Narrative Reportative</td>
<td>Inferentive, Reportative</td>
</tr>
<tr>
<td>Sequential</td>
<td>Habitual</td>
<td>Aorist</td>
<td>Perfect + Presens</td>
<td>-s, (-γūs)</td>
</tr>
<tr>
<td>Personal narratives</td>
<td>Introductive</td>
<td>Preterit</td>
<td>Preterit</td>
<td>-γūl</td>
</tr>
<tr>
<td>Sequential</td>
<td>Aorist</td>
<td>Preterit</td>
<td>Perfect + Presens</td>
<td>-γūl</td>
</tr>
</tbody>
</table>

I believe that this threefold opposition of narrative strategies in these four idioms is a striking areal feature. This confirms the contacts between the Northern Selkup and Nganasan, on the one hand, and between the Northern and Tym Selkup and Vakh Khanty, on the other. It is also obvious that the similarity in the repertoire of narrative strategies can only reflect very “fresh” contact situations: the discourse level is one of the most unstable levels: for example, the well-known borrowing hierarchy proposed in [Thomason, Kaufman 1988] shows that this level is the most permeable to contact influence. This means that narrative strategies (especially their typologically unique repertoire) can hardly remain unchanged for a very long time after the contacts were interrupted. This confirms that relatively recently Nganasan and Northern Selkup still were neighbors on the linguistic map of Western Siberia.