

*Janne Saarikivi (Helsinki, Finland)*

***On the Uralic Substrate Toponymy of Arkhangelsk Region:  
Problems of Research Methodology and Ethnohistorical  
Interpretation***

**1. General**

***1.1. The geographical characteristics of the research area***

European North Russia is probably one of the most thoroughly studied areas with a substrate toponymy in the world. Quite naturally, most of the studies concerning it have been published in Russia and in Russian. Therefore, they may have been left unnoticed by many western scholars. Nevertheless, the study of northern Russian substrate nomenclature is of importance for both the history of Uralic languages and the spread of various groups of Uralic peoples, as well as for the mechanisms and chronologies of the Slavification of the northern Russia.

This article provides an overview of the Uralic (Finno-Ugrian)<sup>1</sup> substrate toponymy of the Arkhangelsk Region (Ru. *Архангельская область*). It serves as an introduction to this research field both for Uralicists and Slavists. It also offers a methodological discussion of the possibilities and limitations of the study of substrate toponyms as well as the problems connected with an ethnic interpretation of northern Russian place names. In this connection some new views which deviate from main-stream Russian research are put forward. Throughout the article, special reference is made to the toponymy of the Pinega basin (a tributary of the Northern Dvina), both because field-work has been carried out in this area by the author and because the toponymy of the area well characterises several general features and interpretation problems of northern Russian substrate toponymy.<sup>2</sup>

The Arkhangelsk Region (320.000 km<sup>2</sup>, 1.336.000 inhabitants) is nowadays an overwhelmingly Russian-speaking region. There are various areas with a Uralic speaking population in its vicinity, however: in the west there is the

---

<sup>1</sup> In this article, Finno-Ugrian and Uralic are used as synonyms. Traditionally, the notion Uralic is used of seven Finno-Ugrian branches and the Samoyed languages. In this article, only toponymy from Finno-Ugrian branches will be taken into consideration.

<sup>2</sup> A monograph by the author of this article on the toponymy of the Pinega District will hopefully be published in the near future. This will provide a larger material basis for the methodological discussions presented in this article.



Republic of Karelia and the Leningrad Region with an indigenous Finnic (Karelian and Veps) population, in the east the Republic of Komi with a Permian (Komi) population and in the northeast the Nenets Autonomous District with Samoyed (Nenets) population. Administratively, the Nenets autonomous district is part of the Arkhangelsk Region, but it stands apart from it in geographical, historical and linguistic respects alike. At its southern edge the territory borders on the overwhelmingly Russian-speaking Vologda and Kirov Regions. To the north the area borders on the White Sea and the Arctic Ocean, but the Kola Peninsula with Sámi (and as a result of 19th century migrations e Nenets and Komi) population is only 60 km away by water.

Most of the Arkhangelsk Region belongs to the Northern Dvina drainage area. To the west, part of the area belongs to the basins of the River Onega and to the north-east to the basins of the Kuloj and Mezen'. All these rivers flow to the Arctic Ocean and the old dwellings in the area are typically situated along them. The landscape is relatively flat. The climate is mostly cold and dry and most of the area is taiga with coniferous forest and marshland. In the extreme north-east the dominant vegetation type is that of the tundra.

At the beginning of the 20th century, the Russian peasant population practiced cattle breeding based on the exploitation of flood meadows and agriculture based on the slash-and-burn method. In addition, hunting, gathering of berries and mushrooms, and, in the north, peasant reindeer herding were practiced. During the 20th century the population has grown rapidly due to industrialisation, the establishment of military bases and, during the Stalin era, due to numerous GULAG prison camps. Simultaneously, forestry has become an important means of livelihood.

In the 1970s it became Soviet policy to abolish the small collective farms. Thousands of villages were declared “perspectiveless” and their inhabitants moved to bigger settlements. This meant considerable changes in the use of the land and in the cultural landscape. After the collapse of the Soviet Union, most of the collective farms have ceased functioning and the concentration of people into larger settlements has been accelerated. These changes threaten to destroy the remnants of the North Russian peasant way of life, which until now has preserved substrate toponyms from the period preceding the Slavic era.

### ***1.2. The present language forms of the region***

The Russian dialects of the area have a twofold historical background. The dialects of the northern part of the region derive from the Old Novgorod vernacular (*древненовгородское наречие, древненоовгородский диалект*) represented in the Novgorod birch bark letters (cf. ZALIZNYAK 2004), whereas



the dialects of the region's southern border are descendants of the central Russian dialects spoken in the upper course of the Volga (cf. KOMYAGINA 1994: 228–232). This state of affairs reflects the twofold origin of the Slavic settlement in northern Russia. The northern and western parts of the Arkhangelsk Region were until 1471 a part of the Principality of Novgorod with the southern and eastern parts being subject to colonisation from the principalities of the Russian central plain—Rostov, Suz'dal, Jaroslavl, Vladimir and—in the later period—Moscow (cf. NASONOV 1951).

The division of North Russian into dialects is quite controversial and is not discussed here. It is enough to note that from the point of view of Russian dialectology, the Arkhangelsk dialects are quite conservative. Most of them share full okanje (non-reduced pronunciation of non-stressed vowels) and cokanje (the merging of two east Slavic affricates into one). They have also preserved *g* between vowels, in a position in which the Russian literary language has *γ* or *v*. Some North Russian dialects also have a glide *v* before a rounded vowel in the word beginning (cf. *вострый* < *острый* 'sharp') and they represent the development *l* > *w* in a postvocalic position in a closed syllable (cf. *доугий* < *долгий* 'long'). Uralic, mostly Finnic and, to a lesser extent, Permian substrate interference is discernable in the vocabulary and in some features of prosody and morphosyntax. The scope and amount of these substrate interferences has been subject to debate for decades and there is no unanimity as to what extent certain North Russian dialectal features, such as the nominative object, the postponed article, changes in accentuation, dialectal merger of voiced and voiceless stops, comparative forms of nouns, etc. have come about due to Finno-Ugrian influence.<sup>3</sup>

As noted above, all the other languages in north-eastern Europe are Uralic. Karelian and Vepsian belong to the Finnic branch of the Uralic languages. These are offsprings of an intermediate protolanguage of the Uralic family, Proto-Finnic. This protolanguage was probably spoken approximately 500 BC–500 AD in the vicinity of the Finnish Gulf (newest dating, KALLIO 2006<sup>4</sup>). The present Finnic settlement of most of inland Finland and Karelia emerged not earlier than the Middle Ages.

The Sámi languages spoken in the Kola Peninsula and northern Fennoscandia (together 25.000–30.000 speakers) are daughter languages of another intermediate Uralic protolanguage, Proto-Sámi. Proto-Sámi has usually been located somewhere in the Onega Region and was probably spoken simulta-

---

<sup>3</sup> Information on proposed phonetic, phonological and morphosyntactic substrate features with references can be found in VEENKER (1967), VOSTRIKOV (1990), SARHIMAA (1995) and SAARIKIVI (2000, 2006).

<sup>4</sup> See p. 2 of the cited article for datings by earlier scholars and also discussion in section 6.5.



neously with Proto-Finnic. Prior to Finnic, the Sámi languages were spoken in most of Finland and Karelia (cf. T. ITKONEN 1948: 88–107; cf. also article by ANTE AIKIO in this volume). Finnic and Sámi have had considerable mutual contact. According to a traditional view (E. ITKONEN 1966; KORHONEN 1981), these language groups also share a common protolanguage within Uralic. Today, this hypothesis finds less support (ITKONEN 1998; KOIVULEHTO 1999a; SAARIKIVI—GRÜNTAL 2005).

Tundra Nenets, spoken in the Nenets Autonomic District, belongs to the Samoyedic branch of Uralic. The languages of this branch are offsprings of Proto-Samoyed, which likely was an earlier protolanguage than Proto-Finnic. As the proto-Samoyed vocabulary reconstruction of JANHUNEN (1977) includes approximately half as much vocabulary as the reconstruction of Proto-Sámi by LEHTIRANTA (1989) and Proto-Finnic has an even greater common vocabulary (cf. HÄKKINEN 1985). The area in which Proto-Samoyed was spoken is in western Siberia, whereas the Nenets of the Arkhangelsk Region are medieval newcomers.

Komi, spoken in the Komi Republic and in the area of the former Komi-Permyak Autonomic District, belongs to the Permian branch of Uralic languages. Further, the spread of Komi to the north and east is a relatively recent phenomenon which took place not prior to the 13th century. The original homeland of the Komi was in the Vychegda river basin in the south of the Komi Republic, and the Proto-Permian homeland was probably even further to the south, in the Kama Region (LASHUK 1970; BELYKH 1999). The dispersal of Proto-Permian is probably a somewhat later phenomenon than the dispersal of Proto-Finnic, dated approximately 700–800 AD (see BARTENS 2001: 10–13). There are Finnic loanwords in the Permian languages and it is therefore obvious that there have been contacts between these two groups of Finno-Ugrian languages (LYTKIN 1967, HAUSENBERG 1983, SAARIKIVI 2005, see also section 6.5.).

In addition to aforementioned languages, extinct Uralic languages may have been spoken in northern Russia. There are historical sources which mention pre-Slavic tribes without parallels among the present-day Uralic peoples (see below 1.3). This issue is discussed in detail in section 6.5.

It is also plausible that prior to or even simultaneously with the Uralic languages, extinct Palaeo-European languages were spoken in northeastern Europe. There is historical (in medieval Russian sources) and archaeological evidence of a tribe called the *печера*,<sup>5</sup> who seem to have stood culturally



<sup>5</sup> The name of this ethnic group is connected with the name of the River Pechora and is derived from the Russian dialectal variant of *пещера* ‘cave’. According to historical sources, the *печера* lived in the caves at the mouth of the river *Pechora*.

apart from the present northern European populations. These people, who lived in northernmost Europe, may have been referred to as the *sihirtja* in Nenets folklore (cf. LASHUK 1958). In the light of multiple substrate borrowings in Nenets, these people were in all probability linguistically non-Uralic. Moreover, there is vocabulary which is probably of substrate origin in Finnic and especially in the Sámi branches of the Uralic languages (SAARIKIVI 2004a; AIKIO 2004) which suggests contact between these language groups and extinct languages (see section 6.5).<sup>6</sup>

### 1.3. Historical and archaeological sources on Finno-Ugrian populations

There are both Scandinavian (Old Norse) and Slavic literary sources on the Pre-Slavic populations of northern Russia. Certain 11–13th century Scandinavian sources call northern Russia *Bjarmaland* and its inhabitants the *bjarmar*. Some facts on the northern Russian Pre-Slavic population mediated by Scandinavian sources indicate that the people of northern Russia were of europoid appearance, spoke a language close to Sámi and practiced agriculture and cattle breeding. Sagas also contain information showing that the *bjarmar* had constant contact with the Slavic principalities in the 13th century (HAAVIO 1965; JACKSON 1993, MELNIKOVA 1986).

Russian chronicles and hagiographies mention several pre-Slavic populations in the present-day Arkhangelsk Region. The tribe name *заволоцкая чудь* which figures in the Russian Primary Chronicle has traditionally been considered the earliest Russian ethnonym of the Finnic population of the Dvina basin (cf. CASTRÉN 1844; HAAVIO 1965). This name is derived from the word *волок* which has been used as a designation for those places where boats were carried over land from one water system to another. In later centuries, the notion *Заволочье* was used of that part of the Principality of Novgorod which was situated in the Dvina basin, outside the basic administrative units, the *пятины* (the ‘fifths’). The other component of the ethnonym, *чудь* has been used of several Finnic tribes in the vicinity of the Gulf of Finland (on the use of the ethnonym see GRÜNTAL 1997; on the origins and use GRÜNTAL *ibid.* and KOIVULEHTO 1997). Besides *заволоцкая чудь* Middle Age Russian sources repeatedly mention the “common” *чудь* in northern Russia. As noted above, there are also historical sources which mention pre-Slavic tribes with no parallels among present northern European populations: *сура поганая*, *тоймичи погане*, *пинежане*, *важане*, *белозерци*, etc. Most of these ethnonyms derive from river names and it has been proposed that at least some of them refer to mixed Slavic-Uralic

<sup>6</sup> In Sámi, the Palaeo-European substrate is, in any case, stronger and newer than in Finnic, where the existence of such a substrate layer may well be questioned.



populations (BERNSHTAM 1973). Some have been clearly hostile to the Slavs, however, as the Chronicles report armed conflicts of Novgorodians with *сура поганая* and *тоймичи погане* during the Middle Ages.

According to the *уставная грамота князя Святослава Олговича*, a historical document written at the time Novgorod emerged as a sovereign principality, northern parts of the Dvina basin were under Novgorod rule even in 1137 (cf. NASONOV 1951; МАКАРОВ 1997: 18–20). By that time, the population was certainly overwhelmingly non-Slavic. This is also reflected in 11–14th century archaeological findings which point to various local groups of Finno-Ugrians. Material culture among some of them (Vaga basin, individual findings in the Pinega basin) shows similarities with the area populated by the Finnic tribes while some (Kokshen'ga and Sukhona basin) had intensive contacts with the Upper Volga region and its pre-Slavic settlers (OVSYANNIKOV 1978; RYABININ 1997; KOLPAKOV—RYABTSEVA 1994). Western influences from the Ladoga region are noticeable even in medieval archaeological findings in the west of the Komi Republic (SAVELJEVA 1992), whereas those findings related to Permian tribes in the Arkhangelsk Region are clearly less important.

During the Middle Ages, Novgorod and the central Russian principalities rivaled over the control of the northern peripheries and their resources. The first Slavs in this region were likely tax collectors and fur traders, who were followed by peasant migrants, probably from the beginning of the 14th century. The Slavicisation of the area was accompanied by an influx of new people from areas where Finnic languages were spoken, localities that were already subject to Novgorod rule. Thus, the migration waves to this area cannot easily be divided into Slavic and non-Slavic. This is emphasised by МАКАРОВ (1997), who has investigated the development of the trade and communication routes connecting the Dvina basin to Slavic centres by analysing archaeological findings from the major watersheds of the Russian European north. In the 12th century, most of these had both Finno-Ugrian and Slavic components. In subsequent centuries, findings connected with the Slavs increased, which seems to point to cultural assimilation of the local Finno-Ugrian populations with the Slavs.

The Slavicisation of the Arkhangelsk Region seems thus to have occurred both by Slavic migration and by a language shift of the Finno-Ugrian population. The latter has consisted of several groups, some of which participated in the same population waves as the Slavs within the Russian principalities which colonised the northern European peripheries. The final linguistic assimilation seems to have taken place in the Late Middle Ages, in some places probably as late as the 16th or even 17th century. The population sta-



tistics continued to have separate entries for Russians and *чюдь*<sup>7</sup> up to the 19th century, however, and even up to the present day there are some barefoot Russians that consider themselves either as Chudes or as the offsprings of the Chudes.

## **2. Toponym systems in northern Russia**

### ***2.1 History of the study of northern Russian toponyms***<sup>8</sup>

Even prior to the first toponymic studies, Finnish and Swedish scholars such as von Becker, Arwidsson and Porthan were aware of the fact that people linguistically close to Finns had previously lived in an area that subsequently became Slavic. This conclusion was inevitable on the basis of Scandinavian sagas and medieval Russian literature. Many historians also demonstrated that there is a rich northern Russian oral tradition concerning the pre-Slavic people of the region (see below 2.2.).

Probably the first linguist to treat the problem of northern Russian toponymy was A. H. VOSTOKOV (1812) who focused on the recurring final components of many northern and central Russian river names. He concluded that these had originated in extinct languages and were remnants of geographical terms. After him, the fennougrist A. J. SJÖGREN (1832a, 1832b) dealt with northern Russian toponymy in several articles dedicated to determining the origin of the Finnic tribes and describing the Uralic peoples. Also, the founder of modern Finno-Ugrian studies M. A. CASTRÉN wrote a small article on northern Russian toponymy (CASTRÉN 1844). He was the first one to point out that, in addition to the Finnish, some toponyms were etymologisable on the basis of the Sámi vocabulary. Some of CASTRÉN's Sámi toponymic etymologies were later mentioned by K. B. WIKLUND (1911) in his treatise on the history of Sámi settlement. Minor treatises on Finnic toponymy in northern Russia were also written in the 19th century by AUGUST AHLQVIST (1887) and MIHKEL VESKE (1890).

The first scholar to systematically collect toponymic material from various sources and interpret the distribution of toponymic types as proof of the pre-historic spreading of languages in northern Europe was D. E. D. EUROPAEUS (1868–70). Quite erroneously, however, he assumed that many central hydronyms of northern Russia and Finland were of Khanty origin. The later work of Europaeus on Ob-Ugrian toponymy was continued by ART-

---

<sup>7</sup> This is the most common ethnonym of the pre-Slavic populations of North Russia (see below 2.2.).

<sup>8</sup> The history of the study of northern Russian toponyms is most thoroughly treated by MATVEEV (2001: 13–47). The sketch presented here relies heavily on this source.



TURI KANNISTO (1927) who asserted that the western boundary of Ob-Ugrian toponyms was much farther east, in the Dvina basin. Even KANNISTO's views were later rejected by MATVEEV (2001) who concluded that there is no convincing evidence of Ob-Ugrian toponyms in the Russian north.

During the first half of the 20th century, eminent slavist MAX VASMER (1934–36, 1941) made an attempt to draw the approximate ethnic boundaries of pre-Slavic Russia on the basis of place names. He used only macro-toponymy and, being ignorant of Uralic historical phonology, made haphazard comparisons based on first-view impressions of the similarity of Russian toponyms and words of Uralic languages. Although he also implemented modern methods, such as a search for parallels of substrate place names in living languages, his results were no more reliable than those of his predecessors. Another eminent slavist JALO KALIMA made interesting remarks on the structure and adaptation of place names such as the observation that the Finnic *s* is substituted both with Russian *s* (*c*) and *š* (*u*) in northern Russian substrate toponyms (cf. KALIMA 1944a, see also KALIMA 1944b, 1946). Regrettably, he did not continue his studies on this topic.

In the Soviet Union of the 1950s and 1960s the Leningrad scholar A. I. POPOV published several articles on the toponymy of Finno-Ugrian origin. He implemented modern methods such as semantic argumentation that referred to those geographical characteristics of the object denoted by the name and took into consideration the role of personal names in toponym formation (for example, POPOV 1965). From the beginning of the 1960s the Sverdlovsk (later Yekaterinburg) scholar A. K. MATVEEV began collecting northern Russian microtoponyms by engaging in fieldwork. MATVEEV and his pupils (most notably M. L. GUSELNIKOVA, N. V. KABININA, V. O. VOSTRIKOV, L. A. SUBBOTINA and O. A. TEUSH) have treated the Finno-Ugrian substrate toponyms of the Arkhangelsk Region in numerous dissertations and articles. As a result, the most common types of northern Russian substrate toponymy have by today been described and provisionally analysed.

According to MATVEEV (1980, 2001, 2004), the main pre-Slavic toponymic layers of the Dvina basin are of Finnic and Sámi origin. It has also been clarified that Permian traces in the toponymy are not numerous and that they are concentrated in the eastern periphery of the region (MATVEEV *ibid.*; 1999). Substantial parallels between the toponyms of southern parts of the Dvina basin and the area historically inhabited by the Merya (Ru. *мерья*), a Central Russian tribe mentioned several times in Chronicles, have also been demonstrated (MATVEEV 1996, 1998).<sup>9</sup> Many interpretation problems con-



<sup>9</sup> The views by MATVEEV concerning the toponyms of this territory have been criticised by AHLQVIST (1997, 2000).

cerning the non-Finnic and non-Permian layers of substrate toponymy remain, however. In addition to Sámi, these layers are referred to as Meryan and *севернофинская* ('North Finnic') by MATVEEV (see discussion in section 6.).

An important contribution to the study of northern Russian toponyms has been made by the Petrozavodsk scholar IRMA MULLONEN. She has studied Finnic and Sámi substrate toponyms along the Finnic-Slavic language boundary in Karelia and adjacent territories (MULLONEN 1988, 1994, 2002). Her studies are based on the simultaneous investigation of living Finnic and substrate toponyms and have yielded reliable results revealing a detailed picture of ancient language contact situations. One should also mention G. Y. SIMINA (1980) and A. L. SHILOV (cf. SHILOV 1999), who have made many interesting remarks concerning substrate toponyms in North Russia.

## **2.2. Russian ethnotoponyms**

In addition to the substrate toponyms, some toponymic models of Slavic origin include information on the pre-Slavic settlers of North Russia. These are mainly ethnotoponyms, which point to contacts between Slavs and other ethnic groups in the area.<sup>10</sup>

The most common ethnonym in the place names of the Arkhangelsk Region is *чудь*. The wide distribution of this ethnonym in place names does not necessarily mean that the Russian European north was ethnically homogenous by the time of the arrival of the Slavs. Most likely, *чудь* was used as a designation for various Finnic tribes. As noted above, a rich tradition of oral history is connected with the Chudes. According to this, the Chudes were white-haired and white-eyed people, who practised cattle breeding and agriculture. When the Novgorodians arrived, the Chudes refused to convert to Christianity. According to legends, the Chudes either buried themselves under the hummocks or moved to "other rivers". These legends also contain information showing that some of the Chudes assimilated to become Russians (PIMENOV 1965; BULATOV 1993). In addition to Russians, the Komis also have similar legends about the Chudes. In the oral tradition of the Sámi, a legendary tribe whose name is etymologically connected to the Russian *чудь*, the *čuh̄ti* (:čud̄i-) are characterised differently to the Russian and Komi traditions concerning the *чудь*, as a hostile and violent tribe (cf. T. I. ITKONEN 1948: 537–545).

The Sámi, Komi and Russian traditions concerning the Chudes and *čuh̄ti* have likely arisen independently. The fact is that some northern Russians

---

<sup>10</sup> A survey of the distribution of ethnotoponyms in the research area was documented has been by POPOVA (1999). The following rests mainly on this source.



have until these days considered themselves offspring of the Chudes<sup>11</sup>, and that the same ethnonym has been used as a self-designation by a group of Finnic people, the Veps. This suggests that *чудь* was probably an endonym of some northern Russian substrate populations.

Other Uralic ethnyonyms have a more restricted distribution in toponyms. Toponyms derived from the ethnonym *Корела* (former *Корьла*) ‘Karelian’ form a couple of clusters in the lower reaches of the Dvina, Pinega and Onega. Toponyms derived from the Nenets ethnonym *самоед* form clusters in the lower reaches of the Mezen’, Pinega, Dvina, Onega and even in the extreme southwest of the Dvina basin. There is also a historical record and oral tradition on Nenets in some present-day Russian parts of the Arkhangelsk Region, such as the mouth of the Dvina (cf. KABININA 1997). The origin of the *самоед*-ethnotoponyms in the south of Arkhangelsk Region remains an enigma. They may be connected with individual settlers, or have a motivation not connected to the Nenets.

The ethnonym of the Sámi, *лопарь*, is present in a few toponyms of the Arkhangelsk Region (see MATVEEV 2004: 192). Even their interpretation is not unambiguous, because the Russians have also referred to the Nenets as the *лопарь*.<sup>12</sup> In addition, there are several dozens of substrate toponyms derived from the stem *лан-*, that is probably related to Finnic ethnonym for the Sámi (Fi. *lappi*). In Finland, ethnotoponyms derived from this stem are commonplace (T. I. ITKONEN 1948: 103). The interpretation of northern Russian *лан*-names is not altogether clear, however. One should note that the ethnonym *lappi* has been also used to refer to Ludes and Karelians (see SAARIKIVI 2004b: 180–181 for discussion).

Ethnotoponyms connected with the Permian people, *зыр(ь)* and *пермь* are found in some eastern areas of the region and, quite surprisingly, also in the basin of the River Ust’ja at the southern edge of the territory. In this area, *зыр(ь)* has also been used as an invective (STE).

### 2.3. Amount, use and systems of substrate toponymy in the Pinega region

Substrate toponyms are common everywhere in the Arkhangelsk Region. Altogether, there must be tens of thousands of primary substrate toponyms in this area (see MATVEEV 2001: 51). Quite naturally, however, the amount and density of substrate toponyms varies according to district. In the Pinega

<sup>11</sup> The author of this article has encountered one man in the village Chakola village of the Pinega District who insisted that he is not Russian but a Chud. This was also confirmed by his neighbours.

<sup>12</sup> In the dialect of Pinega this is the normal meaning of the word. This state of affairs is a further argument for the late appearance of the Nenets in Europe.



District (*Пинежский район*, 41.000 km<sup>2</sup>, 31.000 inhabitants) there are approx. 1200 primary and at least as many secondary substrate toponyms, which is probably around 4–5% of all toponyms (cf. statistics by SIMINA 1980). In hydronyms, substrate toponyms are more common than Slavic names. The flood meadows situated at the bends in the rivers often have names of substrate origin as well. In cultivation names the substrate toponyms are much less commonplace and many of the existing substrate toponyms were probably connected to geographical rather than agricultural objects in the substrate languages. Also, surprisingly many microtoponyms, such as names of meadows, fields and parts of villages are of substrate origin. In addition, there are surnames, nicknames and invectives of likely Finno-Ugrian origin.

As the oldest layer of toponymy, most of the substrate toponyms are macrotoponyms. From these a substantial amount of Russian microtoponyms has been derived. Thus, the river name *Шарда* denotes a tributary of Pinega (in middle course). The name of the river has apparently served as a base for a group of names even in the substrate language, since there is a village *Шардомень* (variants: *Шардонемь*, *Шардоменя*, etc.) at the mouth of the river. This originated from a name connected with the bend of the river (< Finnic \**неети*, see below section 5.1.). Several Russian microtoponyms have been derived from these two macrotoponyms: *Верхняя Шарда* and *Нижняя Шарда* (Upper and Lower *Шарда* river names), *Шардоменский ручей* (brook),<sup>13</sup> etc.

Substrate and Russian toponyms often have the same motivations. In some cases, toponymic pairs of substrate and Slavic names may be interpreted as Russian translations of a substrate toponym (see section 3.2. below). In other cases it seems that Russian and substrate toponyms have been based on the same naming motivation because it has been a natural choice in the context where the names appear. Thus, two brooks named *Нижний* ('lower') and *Верхний* ('upper') *Петручей* presumably derive from the Finnic \**petäjä* 'pine'. These brooks flow into Lake *Сояльское* through pine woods named the *Бор*, a standard North Russian toponym based on an appellative meaning 'pine woods'. This Russian name is probably not connected to substrate names etymologically, but the connected motivations of the names nevertheless support the proposed toponymic etymology for *Петручей*.

In the Pinega region, settlement names of substrate origin are also commonplace. The northern Russian village typically consists of a lengthy chain of small settlements by a river. Typically, the whole chain and its oldest parts

<sup>13</sup> *Ручей* means 'brook'.

have substrate names, while most of the parts have Slavic names. Thus, the oldest part of the village *Лохново* is called *Худгора*, a name connected with the Finnic word stem (Finnish form given) *hiisi* (:hiite-) (in modern language) ‘troll; evil spirit’, (originally:) ‘a sanctuary, centre of a settlement’ (cf. section 5.2.). The second component of the name, *-гора*, is a Russian word meaning ‘hill’, but it has developed to become a sort of settlement suffix in the Pinega dialect. The conclusion that *Худгора* is an old centre of a village can be further supported by the fact that the neighbouring part of the village is called *Усигорка* (< Finnic *\*uusī* ‘new’).<sup>14</sup> Other parts of the village have Russian names.

Many settlement names include elements which, even originally, have been connected to permanent settlements. Thus, the suffix *-ла* typically attached to settlement names in the Finnic languages (*-la*, *-lä*) or the word final name component *-пала* (< *\*palva* ‘village’ see section 5.1.) are commonplace in Pinega settlement names. The fact that many hydronyms are derived from the names of dwelling places (*Военала* village > *Военалка* river, etc.) and that many of the settlement names are etymologisable on the basis of Finnic personal names also points to a surprisingly old age for many settlements.

Quite naturally, there are substantial differences in the distribution of substrate toponyms between villages. These differences can sometimes be interpreted as the result of dissimilar Slavicisation processes. Thus, in the group of villages situated by the River Sura there are especially many (approx. 80) substrate toponyms. Also, a remarkable percentage of the microtoponyms is of substrate origin. It is thus astonishing that over by the River Pinega only a few kilometres away, in the villages of Gorodetsk and Ostrov, just a couple of isolated substrate toponyms are attested. However, a considerable amount of oral tradition on the Chudes has been recorded in these two villages while, in turn, legends of this kind are less characteristic in the villages beside the river. The oral tradition related to the Chudes in Gorodetsk and Ostrov differs from that of many other villages in that it contains legends about warfare between the Chudes and the Novgorodians. There are also historical accounts of the conflicts between the Russians and the “heathens of Sura” (*Сура поганая*) in the 14th and 15th centuries. In connection with this correlation a question arises: could the small number of

<sup>14</sup> According to an old literary source (МИКНР, p. 93) this part of the village has also been called *Новинка* (< Russian *новый* ‘new’). This name can be considered a loan translation of the substrate name. The same source also mentions a parallel name *Чюдикса*, which is connected with the ethnonym *чудь* and serves as a further argument for the pre-Slavic origin of this settlement (DENIS KUZ’MIN, personal communication).



substrate toponyms in Gorodetsk and Ostrov and the simultaneous abundance of oral tradition on the *чудь* be interpreted to mean that these villages were originally founded by Russian newcomers who created their own toponyms and encountered a pre-Slavic population mainly in conflict situations? The villages by the River Sura could then be interpreted as settlements of Uralic language shifters, who preserved their old place names through a language shift. This line of reasoning is further supported by the fact that Gorodetsk and Ostrov are Slavic oikonyms, whereas many old dwelling-place names in the vicinity are of substrate origin.

The above examples demonstrate how the substrate toponyms function together with the Russian toponyms in a network comprising much information about the pre-Slavic settlers in the Russian European north. In most cases, this kind of information can only be obtained through fieldwork.

### **3. Some methodological questions concerning the study of substrate toponymy<sup>15</sup>**

#### ***3.1. The semantics of a toponym as an object for etymological study***

From the point of view of historical phonology, the methods applied to the etymological study of toponyms are mainly similar to the standard methods of historical-comparative linguistics and, therefore, they are not presented here.

One should note, however, that there are some minor peculiarities in the phonological development of substrate toponyms. For example, phonological reduction and dissimilations are more common in toponyms than in the appellative vocabulary and there is more phonological and morphological variation in substrate toponyms than in appellatives. Moreover, unintelligible toponyms maybe subject to folk etymological interpretation. Toponyms with the same lexical content borrowed from a substrate language thus often occur in numerous, slightly different phonological forms in different areas (cf. the Finnish *Kukasjärvi*, *Kuukasjärvi*, *Kuukka*, etc., which all originate

---

<sup>15</sup> As is apparent from the aforementioned, at the present there is an established scholarly tradition in the etymological research of the Uralic substrate toponyms of northern and central Russia. The main references for the methods of such studies are MATVEEV (1986, 2001), GLINSKIKH (1983) and MULLONEN (2002). What follows rests mainly on these sources. Such standard methods as the checking of the old forms of the toponym in the written sources available are left aside here. This method, though useful and important, has severe limitations in northern Russia where the majority of substrate toponyms does not figure in with any early documents.



from the Sámi, cf. North Sámi *Guhkesjávri* ‘long lake’ [see ANTE AIKIO’s article in this volume, cf. also AGEEVA 1989: 94]].<sup>16</sup>

All these peculiarities of phonological development are related to the fact that toponyms may lose their connection with the lexemes they are derived from. This is because the main meaning of the toponym is its denotation (in other words: a place) and not its lexical content (see in detail AINIALA 1997: 15–22). Thus, one of the basic criteria for etymological research, looking for related meanings in the source and target language of the language contact, is not applicable to the study of toponyms.

Though secondary from the point of view of their primary function, all toponyms have a lexical content when they emerge. The formation of toponyms is connected to naming models, which in turn are based on syntactic construction types and lexical conventions (for further references see KIVINIEMI 1977). This means that the same structure, the same lexemes and the same naming motivations recur in thousands of toponyms. This considerably simplifies the identification of lexemes in the case of unintelligible toponyms.

The main methods of the study of the semantics of substrate toponyms are the following:

- 1) Comparative study of the structural and semantic typology of toponyms in substrate languages or languages related to them, the aim of which is to determine common naming models and motivations.
- 2) Study of the geographical characteristics of the objects denoted by substrate toponyms, and checking to confirm that they correspond to the naming models and motivations in the assumed substrate languages.

The successful comparative study of toponyms usually requires place name material not only from the area under investigation, but also from the assumed substrate languages. In many cases we are not able to tell exactly which language this was and, therefore, are forced to use material from related languages.

The perspectives for comparative toponymic studies of Uralic substrate toponymy are relatively good, as many of the Uralic languages have been studied from the point of view of place name typology. In the Uralic languages toponyms are typically compounds consisting of two parts, a specific and a generic. The latter expresses the type of object denoted, whereas the former specifies or qualifies the object by describing those characteristics



<sup>16</sup> Some of the toponyms which derive from the Sámi *guhkes* have folk etymologically been connected with the Finnic *kukka* ‘flower’.

which differentiate it from other objects of the same kind (e.g. Finnish *Kivi/niemi* ‘rock/cape’, literally ‘cape by a rock or with rocky terrain’, *Suoljärvi* ‘marsh/lake’, *Uusi/pelto* ‘new/field’, etc.). The generic is typically a geographical term whereas the specific can be a noun, an adjective or a semantically opaque element. There are also other structural types of toponymy in the Uralic languages such as toponyms derived from participles of verbal stems in Finnic, toponyms formed from action forms of the verbs in Sámi and toponyms formed with a derivational suffix in several Uralic languages. It seems that in language communities with a greater need for toponyms such as the Sámi and Ob-Ugrian communities, which practise a nomadic way of life and occupy large areas of land, deverbal structure types semantically connected with events tend to be more common than in those communities which use only an average number of toponyms. These in turn, use predominantly denominal toponyms connected with the characteristics of the object. In communities with a greater need for toponyms there also seems to be a tendency to create toponyms which consist of more than two lexemes and toponymic clusters consisting of a large number of toponyms.<sup>17</sup> Deverbal substrate toponyms or substrate toponyms consisting of more than two lexemes are not common in northern Russia, however, and this suggests that the Uralic substrate population lived in permanent settlements.

In the course of typological studies of the toponymy of the Uralic languages, the most typical generics and specifics of Finnish (KIVINIEMI 1990), Veps (MULLONEN 1994), Estonian (Saaremaa and Läänemaa dialects, KALLASMAA 2000, 2003), South Estonian (FASTER—SAAR 2001) and Inari Sámi (S. AIKIO 2003) toponymy have already been clarified and similar information is readily available also on Udmurt (ATAMANOV 1988), Komi (TURKIN 1989), Mari (GALKIN—VORONTOVA 2002) and Khanty (DMITRIEVA 2006) toponymy. This information can be used in identifying the recurring elements of northern Russian toponyms.

Thus, for example, the hundreds of toponyms in northern Russia with the seemingly arbitrary final components *-нема*, *-мень*, *-минь*, *-нема*, *-мена*, *-мина*, etc. refer to capes, riversides, and coastal objects. In view of the toponym formation of the Uralic languages, it is obvious that these phonemic chains have originated from a geographical term, more precisely, one that was related to the Finnish *niemi* ‘promontory’ (< Proto-Finnic *\*neemi*).

<sup>17</sup> The observations concerning interdependency between toponymic types and the size of the toponymic system were made by the author when comparing the remarkably different toponym systems of the reindeer Sámi and the Sea Sámi. It seems to find support in the toponymic system of other Finno-Ugrians practising nomadism such as the Khantys (DMITRIEVA 2006). It is the aim of the author to consider this subject in a future publication.



This word is among the most common generics in most of the Finnic languages (KIVINIEMI, MULLONEN, KALLASMAA, FASTER op.cit.). The metathetic forms (-мень, etc.) are explicable in the light of the tendency of Russian to avoid words with a final *-m* while final *-n* is commonplace (MATVEEV 2004: 205).

In a similar manner, hundreds of substrate toponyms in a wide area with the final components *-oï*, *-aiï*, *-oя*, *-ая*, etc. denote brooks. Most of these, quite certainly, originated in Finnic or related Uralic toponyms with the generic *\*woja* ‘brook’ (> Finnish *oja*). This word also belongs to the most common generics in all of the Finnic languages. A related generic is also to be found in Sámi (saN *oadji* ‘brook’ SaK *vuäjj* ‘brook’<sup>18</sup>). In addition, toponyms suggest that a related word has existed even in Udmurtian (ATAMANOV 1988: 61–62).

In addition to generics, the commonly recurring specifics of the substrate toponyms can also be identified on the basis of the living Finno-Ugrian languages. Thus, for example, the Russian toponyms *Кузонемь*, *Явронемь*, and *Котонемь* can be compared with the Finnic (only Finnish forms given) toponyms *\*Kuusiniemi*, *\*Järviemi* and *\*Kotaniemi* (from the appellatives *kuusi* ‘spruce’, *järvi* ‘lake’, *kota* ‘hut; tent’). The specifics of these names belong to those most common in Finnic toponyms. The proposed etymologies are further supported by the fact that these specifics recur in a number of other substrate toponyms as well, although with different endings (e.g. *Кузоя* brook, *Явроньга* river, *Котой* brook).

The recurring word final elements, which typically originate in the generics of substrate languages are referred to as *(topo)formants* (*моноформант*) in Russian toponymic literature. The word initial elements of substrate toponyms, in turn, are referred to as *bases* (Ru. *основа*). Both terms are adopted below. This is because the terms specific and generic do not adequately refer to name elements which have lost their lexical and/or morphological nature.

Despite the fact that formants historically often originate in generics and bases in specifics, formant and base are to be understood as primarily synchronic notions. In substrate toponymy, several assimilative changes may namely affect the shape of the individual toponyms and many formants thus occur in positions in which the corresponding generics are not reconstructable in the substrate language. Moreover, many formants are of multiple origins, though from the point of view of the Russian place name system,



<sup>18</sup> On the basis of its restricted distribution in North Sámi dialects, the former word is presumably a borrowing from Finnic (ANTE AIKIO: personal communication).

they all include phonotactic elements which make it possible to understand them as names. Thus, in the terminology of this study, base and formant may be defined as phonotactic types of one-morpheme opaque toponyms. A characteristic feature of the formant is that it often makes it possible to understand the word as a place name, or sometimes as a name denoting a specific kind of place. The bases do not have this characteristic.

### 3.2. Probability and verifiability of toponymic etymologies

Toponymic etymologies can (and should always, if possible) be supported semantically, also. If a place name that presumably includes a substrate language term for ‘lake’ indeed denotes a lake, or an object close to a lake, this substantially adds to the credibility of the etymology. This is the case with most of the toponyms with the ending *-нeмь*, *-мeнь*, *-минь*, etc. which denote promontories and river bends, or toponyms with the ending *-oй*, *-aй*, etc. which typically refer to brooks.

Some toponymic etymologies are not verifiable on the basis of language-external facts, however. As for *Котонeмь* it is impossible to prove whether the promontory denoted by this name has sometimes been used as a temporary settlement without archaeological investigation. As for *Кузонeмь*, these kinds of names denote various promontories and bends in the rivers, alongside some of which spruce grow while beside others they do not. The proposed etymology may still be correct. It may be that the characteristics of the place have changed during the centuries.

It is also possible that the proposed etymology does not indicate the existence of any features in the denoted object which could verify or falsify the etymology. For example, *Сетала*, the name of a part of a village *Валдокурье* may be connected with Finnic *\*setä* ‘uncle’ as proposed by MATVEEV (2004: 67). However, there is nothing in the object itself that could verify or falsify this etymology. We have to look at different kinds of sources (historical documents, other toponyms, etc.) in order to find support for the etymology and even if this kind of search fails, the etymology could still be correct, though somewhat less probable than many other toponymic etymologies.<sup>19</sup>

<sup>19</sup> In case of *Сетала*, MATVEEV (ibidem) has proposed that the nearby toponym *Чучебала*, presumably derived from the Sámi *\*éacē* ‘uncle’ and *\*palva* ‘settlement’ (see below 5.1.) would support this etymology. The toponyms *Сетала* and *Чучебала*, are also used as synonyms in a 16th century document (MATVEEV 2004: 105–106). The hypothesis concerning of Sámi origins for of this toponym is still incorrect because of the formant which clearly is not Sámi and because of the



Another factor that affects the reliability of toponymic etymology is the frequency of toponymic models in languages used as material for comparisons. The toponymic etymologies referred to above are based on the assumption that common toponymic models of present-day languages were also common in the substrate languages to which they are related. While this certainly is likely, it means also that toponyms based on unusual naming motivations cannot be etymologised with the same degree of certainty as those based on frequent motivations.

The probability scale for toponymic etymologies that follows is based on material from the Pinega District and is, quite probably, not generalisable in all contexts. Furthermore, it focuses only on probability problems related to the semantics of the toponyms as the phonological problems regarding toponymic etymologies can, in the most cases, be accounted for in a similar manner to other etymologies. The toponymic etymologies which fulfill the characteristics for group 1 are, in the opinion of the author, most probable, with the probability diminishing down the scale.

1) Toponyms which belong to toponymic types present in living languages with an etymology that can be verified by language-external facts, cf. *Лимозеро* a lake, *Лимручей* a brook < Finnic *lima* ‘slime’ (the objects are characterised by slime crops)<sup>20</sup>, *Лемозеро* a lake, *Лемонала* a village < Finnic *\*leettek* (> Finnish *liete* ‘sludge’, Karelian *liete* ‘fine sand on a shore’, those places denoted as indeed having a sandy bottom and shores), *Солозеро* a lake < Finnic *\*salo* or < Sámi *\*suolōj* < *\*salo(j)* < *\*salaw* ‘island’ (there is an island in the centre of the lake). *\*lima*, *\*leettek* and *\*salo(i)* are all terms widely used in toponym formation in Finnic. *\*salo(i)* is also frequent in Sámi toponyms.

2) Toponyms with semantically well-founded etymologies that can be verified by language-external facts when there is no corresponding toponymic model in living languages, cf. *Кычас* a river, *Кыча* a lake *Кычверетия* a passway through a marsh<sup>21</sup> < Finnic *\*kiccās* (> Finnish *kitsas*) ‘narrow’ All these names denote objects characterised by their narrowness. However, toponyms with a corresponding appellative are rare in Finnic. The same concept is expressed with several other words (Finnish *kapea*, *kaita*, *soukka* and their counterparts in other Finnic languages).

3) Toponyms which belong to toponymic types present in living languages when the places they denote are neutral regarding the proposed etymology,

---

phonological phonetic shape of the base (this is also admitted by MATVEEV himself, *ibid.*). The assumption by MATVEEV, that the toponym *Чучебала*, derives from the same semantics as *Сетала* may still be correct in principle.



<sup>20</sup> *lima* is a Germanic borrowing and etymologically connected with English *slime*.

<sup>21</sup> *Веретия* is a dialect word that means ‘a narrow dry passway through a marsh’.

cf. *Ристимень* bend in a river < Finnic *\*risti* ‘cross’, *\*neemi* ‘promontory’ (toponyms formed from the appellative *risti* ‘cross’ are common in Finnic languages, but there is no evidence that there was any kind of a cross in that place, or that the place would have been situated at some kind of cross-roads), *Ламбас* two brooks, one lake < Finnic *\*lampas* ‘sheep’ (toponyms formed from the appellative *\*lampas* are common in Finnic languages and the objects denoted are relatively close to old dwellings and could thus have been connected to sheep herding. However, this would seem impossible to demonstrate),<sup>22</sup> *Сергозеро* < *\*säarki* ‘roach’ (roach is a common fish in luxuriant lakes of northern Europe and it could be a possible naming motivation for a large number of lakes in any district).

4) Names connected to appellatives not used (or very rarely used) in toponym formation in living languages while the object is neutral in regard to a proposed etymology, cf. *Рачмина*, *Рачканда* < Finnic *\*raccu* ‘mount; riding horse’ (MATVEEV 2001: 63). Etymologies of this kind are extremely uncertain and in many cases probably false.

A fifth group of toponymic etymologies which does not need to be placed in the probability scale is the toponymic etymologies proper, that is, toponyms which may be connected with each other while no appropriate etymological explanation for them can be given. Thus in the Pinega basin there are two rivers called *Кырас*. On phonological criteria, they may be connected with Finnic hydronyms derived from specific *kyrö(s)*-. In Finland, similar names are connected to several rapids and stony places by rivers, or to fast flowing rivers.<sup>23</sup> The element *kyrö(s)* itself, however, is without an etymological explanation. Another example is the river and village name *Турья* that may be connected to several Finnic toponyms with the specific *turja*-. No credible etymological explanations for this have been given. Nevertheless, such correspondencies can point to links between the toponymy of certain regions and thus help to clarify problems related to settlement history.

<sup>22</sup> MATVEEV (2004: 45–47) has proposed a connection of this toponymic model and the Russian dialectal *ламбас* ‘bay of a river’, which is, in turn, almost certainly is a borrowing from a Finnic word related to Finnish *lampi* ‘small lake’. In the Pinega district, those toponyms derived from *\*lampas* are not connected to river bays or small lakes, however, and this makes the etymology proposed above more probable in the given context. Similar toponyms connected with the Finnic *lampi* and the Russian *ламбас* exist in other districts.

<sup>23</sup> There is also a homonymic western Finnish toponymic type *\*kyrö(s)* which is derived from *\*kydö* ‘moorland burnt-over for cultivation’ (> Fi. *kytö* id.). The hydronyms derived from the stem *\*kyrös* have a wider distribution and are not connected with these.





Mordvinian branches of Uralic which all have voiced consonants or Udmurtian which follows word final stress pattern and, in these respects, been similar to most of the Uralic language forms, both modern and reconstructed.

The phonemes of substrate languages are most easily reconstructed in the first syllable and in the consonant cluster between the first and second syllable. The second syllable of the substrate name typically has reduced vowels, and if this syllable is word final, it also includes the Russian gender ending, which is typically determined by the gender of the geographical appellative that characterises the object. Thus, village and river names are typically feminine (and end with an *a*), because the words *деревня* ‘village’ and *река* ‘river’ are feminine, while brook names tend to be masculine (< *ручей* ‘brook’) and lake names neutre (< *озеро* ‘lake’). This substantially diminishes the possibility of some vowels occurring in word final position. Thus, the reconstruction of substrate language phonemes in the second syllable can usually be made only at the lexical level when the word the toponym is derived from has been identified with the aid of the first syllable.

The central sound correspondences of Finnic loanwords in Russian were clarified at the beginning of the 20th century (MIKKOLA 1894; KALIMA 1919; see also MYZNIKOV 2004: 345–371). The sound correspondences found in northern Russian substrate toponyms are mainly close to these. They are not completely uniform in the entire Dvina basin, however (see MATVEEV 2001: 123–151).

The following table includes the most typical sound correspondences of the toponyms in the Pinega district. Some correspondences in other northern Russian areas are discussed in the footnotes. Most of the correspondences included in the table can be supported by several etymologies from groups 1, 2 or 3 in the reliability scale presented above.

**Table 2.**

**A) CONSONANTS**

<i>t</i>	<i>t, d</i>	<i>Торос-</i> < Pre-Sámi <i>*toras-</i> ‘crosswise’, <i>Хум-</i> / <i>Худ-</i> < Finnic <i>*hiisi</i> (Sg. Gen. <i>hiite-</i> , Pl. Gen. <i>hiitte-</i> ) ‘sanctuary; centre of a settlement’
<i>k</i>	<i>k, g</i>	<i>Каск-</i> < Finnic <i>*kaski</i> ‘burnt-over clearing’, <i>Соз-</i> < Finnic <i>*soka</i> ‘dirt; litter’
<i>p</i>	<i>p, b</i>	<i>Палм-</i> < <i>palttV-</i> ‘slope’, <i>Ламбас-</i> < <i>*lampas</i> ‘sheep’
<i>tt</i>	<i>t</i>	<i>Хатар-</i> < <i>*hattara</i> ‘bush’
<i>kk</i>	<i>k</i>	<i>Азук-</i> < <i>*Asikka</i> personal name



<i>pp</i>	<i>p</i>	?? <i>Лyn-</i> < Karelian <i>luppo</i> ‘water lily; Nymphae or Calla palustris’, <sup>24</sup>
<i>č</i>	<i>č</i>	<i>Чолм-</i> < Proto-Sámi <i>*čoalmē</i> or Pre-Finnic <i>*čoľma</i>
<i>c</i>	<i>č</i>	<i>Печ-</i> < Proto-Sámi <i>*peaccē</i> or Pre-Finnic <i>*pecä</i> ‘pine’
<i>m</i>	<i>m, n</i>	<i>Лим-</i> < Finnic <i>*ľima</i> ‘slime’, <i>Ланбас</i> < <i>*lampas</i> ‘sheep’, <sup>25</sup>
<i>ŋ</i>	<i>n</i>	<i>Вонга</i> < <i>*vojka</i> ‘deep place in riverrun’
<i>s</i>	<i>s, š, z, ž</i>	<i>Шул-</i> < <i>sula</i> ‘unfrozen’, <i>Сул-</i> < <i>*sula</i> id. <i>Перз-</i> < <i>*perse</i> ‘back’ (in modern language: ‘rump’) <sup>26</sup>
<i>h</i>	<i>h, g</i>	<i>Хум-</i> < <i>*hiisi</i> ‘sanctuary’ (cf. above, see also section 5.) <i>Карз-</i> < <i>karhi</i> ‘harrow’ (in toponyms also ‘rough’)
<i>hk</i>	<i>ht</i>	<i>Пухт-</i> < <i>*pihka</i> ‘resin’ <sup>27</sup>
<i>Vu</i>	<i>Vv</i>	<i>Лавд-</i> < <i>lauta</i> ‘board’ (in toponyms ‘trap’)

The following substrate language consonants always correspond to parallel consonants in Russian: *l, r, v, n*. The phoneme *j* also corresponds to the vowel prothesis or to *j*.

Russian *p, t* and *k* are regular correspondents of the substrate language *\*p, \*t* and *\*k* in word initial position and next to an unvoiced consonant. Russian *b, d* and *g* are regular correspondents of the substrate language *\*p, \*t* and *\*k* between vowels and next to a voiced consonant. In some cases however, unvoiced *t, p* and *k* also seem to occur in these positions. It is possible that toponyms of this kind originate from derivations. Thus, it seems likely that the brook name *Ремова* (var. *Рѣмова*) is related to the Finnic *\*retu* ‘dirt’ as this word commonly occurs in Finnic toponyms. The Russian voiceless *-t-* hints that it goes back to the substrate language plural stem derivation *\*retto(i)*—this kind of derivation also appears in Finnish and Karelian toponyms (*Rettoinsuo, Rettuisuo, Retteinnotko*, etc. [NA]). The alternation *xum* ~ *xud* ‘sanctuary; centre of settlement’ may also have arisen because toponyms with the base *xum-* originated in plural forms (cf. Finnish *Hiittensuo*, etc.), whereas toponyms with the base *xud-* suggest a singular (cf. Finnish *Hiidenvuori*, etc.).

<sup>24</sup> This etymology is very insecure in that the base *Лyn-* only occurs in the toponym *Лynмега*. It is not clear what the *-pt-* stands for. However, according to informants, *Лynмега* is a river covered by water lilies.

<sup>25</sup> The correspondence *\*m ~ n* is a rare one and clearly is a result of sporadic dissimilation.

<sup>26</sup> It is not clear, how many sibilants there were in substrate languages (see discussion in section 6.3.).

<sup>27</sup> This somewhat surprising correspondence also occurs in some other districts. In most of the dialects, however, *-hk-* has also been substituted as *-hk-* and *-vk-*.

Some occurrences of the phoneme *\*t* in Finnic originate from the phoneme *\*ð*, cf. name of the river *Сотка* < Proto-Finnic *\*sotka* < Pre-Finnic *\*šodka* ‘wild duck’. It is not clear whether this phoneme was preserved in some substrate languages. In any case, its reflexes are the same as those of *\*t*.

As the northern Russian dialects are characterised by *cokanje* (i.e. they have only one affricate) it is impossible to trace back the possible different reflexes of two Finno-Ugrian affricates. Although both affricates occur in the etymons of the substrate toponyms, they represent only one phoneme and it is impossible to know whether the two Finno-Ugrian affricates were present in the substrate language (see section 6.4. for further discussion).

The occurrence of *h* and *g* as reflexes of the substrate language *\*h* depends on the phonological environment. *g* is a regular correspondent of the substrate language *h* next to a back vowel, *h* next to a front vowel.

**Table 3.**

**B) VOWELS**

<i>a</i>	<i>a, o</i>	<i>Варгас</i> < <i>*varkas</i> ‘thief’ <i>Полта</i> < <i>paltte</i> ‘slope’
<i>e</i>	<i>e, ë, a</i>	<i>Кёлд-</i> < <i>*kelta</i> ‘yellow’, <i>Падр-</i> < <i>*petra</i> ‘wild reindeer’
<i>*ee</i>	<i>e</i>	<i>Лет-</i> < <i>*leettek</i> ‘fine sand’
<i>i</i>	<i>i, i̇</i>	<i>Пим-</i> < <i>*pime(ðä)</i> ‘dark’, <i>Кыч-</i> < <i>*kicca(s)</i> ‘narrow’
<i>o, oo</i>	<i>o</i>	<i>Вонга</i> < <i>*voŋka</i> ‘adeep place in the river’ <i>Лод-</i> < <i>*loodeh</i> ‘west or south’
<i>u</i>	<i>u</i>	<i>Руск-</i> < <i>*ruske</i> ‘red or brown’
<i>y</i>	<i>u, i̇</i>	<i>Кыл(ь)м-</i> < <i>*külmä</i> ‘cold’, <i>Юрома</i> < <i>*jyr(h)ämä</i> ‘a river that runs through a lake’
<i>ä</i>	<i>ä, e, a</i>	<i>Хярг-</i> < <i>*härkä</i> ‘bull’, <i>Серг-</i> < <i>*särki</i> ‘roach’, <i>Сейвас</i> < <i>*seiväs</i> ‘(hay) pole’ <sup>28</sup>
<i>ö</i>	<i>?i̇</i>	<i>Выр(ь)-</i> < <i>??*vöörü</i> ‘slope’ <sup>29</sup>
<i>*aj &gt; ej</i>	<i>?aj</i>	<i>Хайн-</i> < <i>??*haina</i> ‘hay’ (> Fi. <i>heinä</i> )

<sup>28</sup> In the Pinega district, the correspondence substrate language *ä* ~ Russian *a* is limited to the second syllable. It is also attested also in first syllable in some other northern Russian territories.

<sup>29</sup> Because the *ö* is an infrequent vowel in Finnic, this correspondence is quite insecure. In Finnic, there are two close words with the meaning ‘slope’ *\*veere* (> Finnish *vieri-*) and *\*vöörü* (> Finnish *vyöry-*). The latter, which only occurs in toponyms, clearly is a labialised variant of the former, but the fact that it is present in both the northern and southern groups of Finnic languages points to its high age. Thus it is well quite possible that the Pinega toponym *Вырполье* which indicates a field situated on a slope in the village of Krylovo may be connected with this Finnic word.



According to the standard interpretation (MATVEEV 2001: 133–136), the correspondence  $a \sim o$  is older than  $a \sim a$ . The Russian short  $a$  developed into  $o$ , but this development may be more recent than generally assumed (JUHANI NUORLUOTO: personal communication based on a new interpretation of the occurrence of vowel graphemes in the Novgorod birch bark letters). The toponymy of the Pinega district supports this hypothesis in that here the correspondences  $a \sim o$  and  $a \sim a$  occur in the same area.

The correspondents of  $i$  and  $e$  are determined by the vowel in following syllable. If the second syllable has a back vowel the regular correspondents are  $i$  (ortographic  $ы$ ) and  $a$ . If the second syllable has a front vowel, the correspondents are  $e$  and  $i$  (see also MATVEEV 2001: 137–138; MATVEEV 2004: 205–210). In southern Finnic, a mid-central vowel (in Estonian orthography  $\ddot{o}$ ) has emerged in the first syllable of the words which have the combination  $e \sim a$  (HOLST 2001). It is not impossible that a similar process might have also occurred in the Finnic substrate languages of the Dvina basin.

#### 4.2. Morphological adaptation

Several morphological adaptation techniques are applied in the integration of substrate toponyms into Russian. At least the following morphological integration types can be distinguished.

1) The full adaptations (Finnic *\*Kuusineemi* ‘fir/promontory’ > Кузомень, *\*Haaraniemi* ‘branch(ing)/promontory’ > Харомень, *\*Kuusioja* ‘fir/brook > Кузоя).

In these cases the Uralic substrate name typically consisting of a generic and a specific has been borrowed into Russian as a single-morpheme name. Thus, Finnic *\*Kuusineemi* is a syntactic construction that consists of two intelligible appellatives, but the Russian Кузомень is an arbitrary one-morpheme name which cannot be segmented in the language in which it functions. Thus, although the formants are word final, from the point of view of morphology, they are more like stem types than suffixes.

As the same formants recur in thousands of toponyms, the relationship between them and the types of objects they denote is often more or less obvious. This may have resulted in a limited consciousness by Russian speakers that, for example, the phoneme chain *-мень* usually denotes a promontory or a bend in a river. This may lead to a kind of “remorphemisation” of the substrate name what can be observed from the fact that sometimes formants develop analogically in Russian from other word-final elements in order to keep the name in shape with language-external facts (cf. *Торома* > *Торомень* > *Торонемь*, as the object denoted is situated on a promontory).



2) Partial translations (\**Limajärvi* > *Лимозеро* ‘slim/lake’, \**Petä(jä)oja* > *Петручей* ‘pine/brook’).

In these cases the generic of the name is translated into Russian while the specific remains untranslated. As a result, the substrate toponym consist of two morphemes, the latter of which is a Russian geographical appellative functioning as a classifier and the former a lexically arbitrary element that carries the denoting function of the name.

The number and types of partly translated names vary according to the type of object and the area. The names of the lakes and marshes tend to be partly translated, whereas the names of rivers hardly ever are, while again, names of brooks are translated in some areas and in some areas they are not (GUSEENIKOVA 1994). In some cases the phonological similarity between the substrate language word and its Russian counterpart may have favoured partial adaptation (?cf. \**vaara* ‘hill’ > Ru. *гора* ‘hill’).

As mentioned in section 3.1., some topoformants, especially those connected to river names (*-га, -н(ь)га, -ма*), can to some extent fulfill the function of a generic also in Russian. This is probably the reason why river names are rarely partial translations.

3) Elliptical toponyms (\**Kylmäjoki* ‘cold/river’ > *Кыл(ь)ма*, \**Lampasjärvi* ‘sheep/lake’ > *Ламбас*).

With this kind of toponym the generic of the substrate language has disappeared and only the original specific of the name functions as a one-morpheme substrate name. In Finnish onomastics, such names have been referred as elliptical.

In some languages (including Finnish), etymologically opaque toponyms have a tendency to shorten by abolishing the generic (cf. Finnish *Kymijoki* > *Kymi*). In the Pinega district the borrowing of a substrate name as an elliptical toponym always occurs when the last syllable of the original specific would have yielded, as a result of phonological adaptation, a syllable identical to the common topoformant. Thus, the river name *Кылма* has a final syllable similar to place names with the formant *-ма* (see below section 5.1.) and this seems to be the reason why the second component of the river name has disappeared. Sometimes, however, the disappearance of the generic is not connected to the phonological form of the name in any way.

Elliptical shortenings seem to be especially common in river names, probably because these are the most important names in the toponym systems of northern Russia and often serve as bases for other names.

4) Suffixations (\**Vihto(j)* personal name > *Вухтово*, \**Kylmäoja* cold/brook > *Кылмовка*).



In these cases, the substrate name has been adopted with the aid of a Russian suffix. In some cases, the suffix has probably replaced a generic or a derivational suffix of a substrate language.

Many settlement names of substrate origin have been formed with the aid of the suffix *-(o/e)vo* which is typically attached to personal names or to toponyms derived from personal names (*\*Aino(i)* personal name > *Айново* village [cf. section 5.3.], *Toivottu* personal name [*< toivottu* past passive participle from *toivo-* ‘hope (v)’] > *\*Toivottula* > *Тойвотолово* village, cf. SAARIKIVI 2003: 140, note 93). In some cases, these kinds of suffixes may have replaced a substrate language derivational suffix *\*-la/-lä* (cf. below section 5.1.), other names for this type (probably a majority of them) may be genuine Russian names derived from Finnic personal names.

In addition, many brook names have been adapted to Russian by attaching a deminutive suffix to the name stem. In other names, suffixation rarely occurs. This is apparently due to the fact that topoformants function in northern Russian dialects in a somewhat similar way to suffixes. As to the latter, they also carry the information that the word belongs to the class of names.

5) Translations (calques) (*\*Seiväsjoki* ‘pole/river’ > *Сейвас* > *Жердь* ‘pole’, *\*Rautaveräjät* > *Железные ворота*).

In these cases the whole name has been translated into Russian. Translations can be identified if the substrate language toponym has been preserved in a literary source, or (and what is more common in northern Russian circumstances) if a substrate toponym with similar lexical content has been preserved in the immediate proximity of the Russian toponym. Thus, the Pinega District river name *Жердь* formed from an appellative with the meaning ‘balk; pole’ and the river *Сейвас* (< Finnic *seiväs* ‘pole’, probably an elliptical toponym from *\*Seiväsjoki*) are situated only one kilometre away from each other. Therefore, it seems quite probable that the Russian name is a translation of the latter. This is further supported by the facts that the Russian name represents a structural type not common in Russian toponymy (the name is composed of a substantive only) and that river names of Slavic origin are otherwise rare in the Pinega district.

Most likely, many translated toponyms will not be identifiable because of a lack of literary documentation and substrate names with a similar lexical content.

6) Full or partial folk etymology (*Лодозеро* river < *\*loodesara* ‘west/brook’, *Рандростров* < *\*Rantasara* ‘shore/brook’).

In these cases the substrate name has been adapted to Russian by mixing it (or a part of it) with a Russian appellative that resembles its phonological shape. The result is an (at least partially) intelligible Russian name that lacks semantic motivation. Thus, *Лодозеро* is seemingly a lake name. The object



it denotes is not a lake, however, and there is no lake in its vicinity. The name denotes a river that forms the upper end of a water system in the basin of the River Pokshen'ga. A characteristic feature of the River *Лодозеро* is that it flows into the Pokshen'ga straight from the west. This would make it possible, although with reservations, to connect the name etymologically with the Finnic *\*loodē(h)* 'west' (in modern Finnish: 'southwest'—this etymology by DENIS KUZ'MIN, personal communication). The Russian second component *озеро* would, in this case, have originated from *\*sara* 'a river at the top of the water system' (see section 5.1.).

In a similar manner, *Рандростров* is apparently an island name. The object it denotes is a brook, however. As all the island names in the Pinega district are partial translations one could, although with reservations, connect this name etymologically with the appellative *\*sara* 'brook' (see below 5.1). In this case, the phonological similarity of *\*sara* and *\*saari* 'island' would have produced an erroneous translation (GUSEENIKOVA 1994). It is even possible that the folk etymological mixing of *\*sara* and *\*saari* has happened in the substrate language and reflects the fact that there were two closely related Finnic substrate languages in the area (see below section 6.3. for discussion).

## **5. Most frequent elements in Russian substrate names**

### ***5.1. Most common formants and their origin***<sup>30</sup>

In what follows some representative toponymic models of north Russian substrate toponymy are presented.

The first list includes the most common formants of the substrate toponymy of the Arkhangelsk Region. As noted above, most of the formants originate from geographic appellatives. Some formants, especially those denoting rivers, seem to be of multiple origins. Thus, those names, which at present include same formants, have not necessarily been of same structure in the substrate languages. This is because in those circumstances in which large amounts of substrate toponymy are borrowed, unintelligible toponyms easily affect the phonological shape of one another. When enough substrate toponyms with similar endings are borrowed, they may turn into a structural toponymic model which, in turn, begins to affect the adaptation of new

---

<sup>30</sup> The most common formants of northern Russian substratum toponymy are presented and etymologically analysed in several articles, manuscripts and a recent monograph by A. K. MATVEEV (MATVEEV 1980; 2001; 2004). The following discussion rests heavily on these sources. In certain cases, however, the views presented below will differ from those of Matveev.



toponyms. There are cases where substrate names analogically adopt new formants in Russian. This kind of reorganization of the toponymic system is a continual process and sometimes there are concurring forms of many toponyms with different formants used simultaneously (*Торома* ~ *Торомень* ~ *Торонемь*, cf. above section 4.2.).

For all the formants below, the following information is given: 1) the most common form of the formant and its main variants in brackets, 2) some examples of toponyms which include the formant, 3) a relative number of toponyms which include the formant in the Arkhangelsk Region (mainly according to MATVEEV 2004), 4) the classes of objects the formant is connected to and 5) the proposed etymology.

**-Вн(ь)га** | *Шиленьга, Покишеньга, Явроньга* | rivers | several hundreds | The formant is of multiple origin. Some names originate in a combination of Uralic genitive *\*-n* and PU *\*juka* ‘river’ or one its successors (as already pointed out by SJÖGREN). Some are analogical formations and have originated in Russian from toponyms with different word final elements. Some names are possibly connected to Finnish toponyms with the suffixes *-nki, -nko, -nka*. Also, this Finnic group is of multiple origin (see RÄISÄNEN 2003), but some of the toponyms in this group are probably of considerable age.<sup>31</sup> Furthermore, the somewhat fantastic etymological suggestion by A. L. SHILOV that toponyms with this formant could include a Uralic word connected to Khanty (Proto-Khanty form given) *\*jeŋk* ‘water’ (< *\*jeŋi*) could also find some support, in that two other common toponymic types (*ухт-, -нала*) are also connected to Uralic words surviving only in the Ugric languages.<sup>32</sup>

**-мень** (*-немь, -нема, -мина*) | *Кузонемь, Шуламень, Каскомень, Чухченема* | several hundreds | villages, capes, river bends, flood meadows, coastal objects | < Finnic *\*neemi* ‘promontory’ (the form *-мень* has come about through metathesis caused by the unusual word final *-мь*). The word *\*neemi* is without cognates outside Finnic and without a generally accepted etymology.

<sup>31</sup> Although RÄISÄNEN has presented etymologies for most of the Finnish toponyms with these endings, some of them are quite dubious (they would belong to groups 3 and 4 on the reliability scale presented in section 4 above). Those Russian toponyms with the formant *\*-н(ь)га* are equally enigmatic. It is possible that among the both groups of names there are pre-Uralic toponyms. This seems likely in that many names of this kind refer to objects of considerable size and even their bases are difficult to etymologise.

<sup>32</sup> A. L. SHILOV further suggests that the Mari *eŋer* ‘river’ with its cognates in especially Central Russian substrate toponymy (MATVEEV 1998) would also belong to this connection as derivations. However, the Mari word derives from Proto-Uralic *\*eŋi-* while the Khanty word points to Proto-Uralic *\*jäŋi* (> Fi. *jää* ‘ice’). Therefore, this explanation cannot be correct.



**-га** (-юга, -юг, -уэ) | Немнюга, Ежуга, Пинега | rivers | approx. 200 | The formant is of multiple origin. Most of the names with this ending, quite certainly, originate from PU *\*juka* ‘river’ and the words related to it (> fi. *joki*, SaN *johka*, Komi *ju*, etc.). Some of the names with this ending originate in words with a derivational suffix (*\*-k*, *\*-kkV*)<sup>33</sup> and some are the result of analogical name formation or adaptation in Russian.

**-(в)ой** (-бой, -буй, -ой, -уй, -оя, -уя) | Каргоя, Кукобой, Мурдой | brooks | < 200 | < PU *\*woja* ‘brook’ (> Fi. *oja*, SaKi *vuäjj*, saN *oadji*). Northern Russian substrate languages clearly had two lexemes related to the Uralic word meaning ‘brook’, *\*oja* and *\*woja*. The latter of these has been characterised as Sámi by MATVEEV (2001) but this is not inevitable because both the Finnic *oja* and the related Sámi words derive from *\*woja*.<sup>34</sup> Therefore, those names which go back to the substrate language *\*woja* can ultimately also derive from another kind of Uralic language than Sámi.

**-ма** | Торома, Мадома, Полтома | rivers, meadows, coastal objects | < 200 | Most of the names with this formant originate from various suffixes of Uralic languages (see discussion by MULLONEN 2002: 222–228). These include deverbial suffixes (most notably *-mA*, deverbial nominal suffix and the suffix *\*-mV* often attached to geographical appellatives (cf. Finnish *oja* ‘brook’, *virta* ‘stream’, *reuna* ‘rim’ > *ojama* ~ *ojamo* [< *oja* ‘brook’], *virtama* ~ *virtamo* [< *virta* ‘stream’], *reunama* [< *reuna* ‘rim’], etc.). The suggestion that toponyms with this ending could have originated from the Uralic *\*mijj* (> Finnic *maa*) ‘earth’ (MATVEEV 2001: 200–202) is, in most of the cases, probably false.<sup>35</sup>

**-саp(a)** (-саpа, -соpа, -зоpа, -зоп, -зур, etc.) | Соросаpа, Лавзоpа, Явзоpа | rivers, brooks, especially the uppermost brooks of the water systems | < 100 | ? < Finnic *\*sa(a)ra* ‘brook, branch of river’. The meaning attested in substrate toponyms is close to another Finnic appellative *haara* (< *\*hara* < *\*šara*) ‘branch’, but the two Finnic words referred to are not etymologically connected (the former is probably a Sámi borrowing (AIKIO 2001), the latter a Baltic loan (cf. Lithuanian *žarà* ‘branch’, JORMA KOIVULEHTO, personal communication with ANTE AIKIO). One should also note that there is no liv-

<sup>33</sup> These kinds of suffixes are common everywhere in Uralic and reconstructable in Proto-Uralic. *-k* is deverbial (cf. Finnish *lähte-* (< *\*läkte-*) ‘to commence; to leave’ > *lähde* (< *\*läktek*) ‘source; spring’. *-kkV* forms collective denominal derivations (Finnish *kuusi* ‘fir’ > *kuusikko* ‘woods that grow fir’).

<sup>34</sup> In Finnic and Sámi, word initial *\*wo* developed into *o* (cf. PU *wolka* ‘elbow’ > Fi. *olka* SaN *oalgi*). East Sámi and also Livonian have a secondary vowel prothesis.

<sup>35</sup> In Finnic, *maa* is used in toponym formations mainly as a part of compounds (*sydänmaa* ‘heartlands’, *palomaa* ‘burnt land’, etc.). There are also some other naming models with the generic *-maa* (‘large island’, etc.).



ing Finnic language with a high frequency of the word *sa(a)ra* in toponyms. It has a limited area of distribution in southeastern Finnish dialects, Veps and Ludian. However, even bases of northern Russian toponyms with the formant *-сара* are often etymologisable on the basis of Finnic languages. This suggests that the language in which the *\*-sar(a)*-names originate was likely different from living Finnic languages.

**-ч** | *Вадасеч, Кокач, Комуч* | brooks, small lakes | approx. 100 | < Karelian *\*-ččU* (a diminutive suffix). Also Sámi has a *č*-diminutive although this is of another origin (< *\*-ńće-*) and some names of this kind may be connected with it.

**-нал(а)** (*-пол(а), -бал(а), -бол(а)*) | *Летопала, Кушкнопала, Военала* | villages, settlements, coastal objects | over 50 | < ?*\*palwa* ‘settlement’ (> Khanty V *рүл*, etc., Mansi TJ *рәл*, etc., Hungarian *falu* ‘village; settlement’). In the present Finnic languages, the word *\*palva* is not used as an appellative, but it has probably been preserved in Estonian toponymy as the component *-palu* in some settlement names.<sup>36</sup> It seems clear that, at least in northern and central Russia, toponyms with this formant denoted settlements even in the substrate language. The comparison with PU *\*palwa* presupposes a somewhat unexpected phonological development in the second syllable, where *\*w* should have disappeared. This development could well have been caused by the adaptation of toponyms into Russian in some dialect, from which the formant would have spread further by analogy. Another possibility is that the second syllable development *va > u* took place in the substrate language.<sup>37</sup>

Some, but likely few names with this formant may have originated from the Finnic *\*palo* ‘burnt land’ and *\*pooli* ‘half; side’, in toponyms also: ‘region’.

**-важ** (*-ваш, -маж, -маш*, etc.) | *Роваж, Игловаж, Косваж* | brooks, rivers | over 50 | < Proto-Permian *\*vož* (> Komi *vož* Udmurt *viž*) ‘branch, brook’. The variants of the formant are explainable on the basis of the phonological environment of the formant. In addition to Permian, there is a word *vož* ‘branch of a river’ with a toponymic use also in Mari, where the word can be considered a borrowing from Permian. The Permian etymology of the formant is verified by the fact that even the bases occurring with this formant are etymologisable on the basis of Permian.

<sup>36</sup> Most of the Estonian toponyms with the with final component *-palu* are, without doubt, connected to the appellative *palu* ‘burnt land’ but in some cases the origin of the names is not altogether clear.

<sup>37</sup> Cf. Estonian *palve* ‘request’ but *palu-da* ‘to request’ (< *\*palvu*), where the derivational suffix *u* has triggered the assimilation *vu > u*.



**-вей** | *Вырвей, Эльвей, Тылвей* | brooks | under 50 | < Proto-Permian \*vVj ‘brook’ (Komi *-vej* in place names; ud. *vaj* ‘branch; brook’ [latter meaning in place names]). As noted by MATVEEV (2001), the Permian character of the names with the formant *-вей* is obvious both on the basis of their distribution and the fact that the bases of the names are usually etymologisable from the Permian languages. One should note, however, that there is a similar word in the Sámi languages as well: saN *veadji* ‘brook’ (< \**vejä*). The Sámi and Permian words cannot be cognates, but the Permian word could be a western Uralic borrowing (see discussion in section 6.4.).

**-ла** | *Веркола, Чакола, Кеврола* | settlements | over 50 | < Finnic *-la*, a locative suffix added to place names. This suffix has developed into a suffix of settlement names exclusively in Finnic, but it has etymological cognates in other Uralic languages.

**-вера** (*-бера*) | *Матвера, Пимбера, Русковера* | settlements, hills, slopes | ?30 | < \**veeri* ‘hill; slope’ > Finnic *vieru, vieri, vieremä* ‘slope’, Proto-Sámi \**vēre* (> saN *vierra* ‘hill on which trees grow’). Also, mdE *ve’* mdM *vär* ‘upwards’ belong here. The semantics of the places denoted by this formant in the Pinega district are similar to that of the Finnic words. Surprisingly, many of these denote settlements, but as the settlements in the Pinega district are typically situated on high places beside rivers, it is not possible to decide which meaning was the original one. Note that in Estonian, a common settlement name model with the ending *-vere*, has most likely developed from \**veeri* ‘slope’ (KETTUNEN 1955: 272–324).

**-вара** (*-вора*) | *Кочевар, Падчевары* | hills | approx. 20 | < SaN *várri* ‘hill’ < PS \**vārē* or Fi. \**vaara* ‘hill’. The Finnish and Karelian *vaara* is, most likely, a borrowing from Sámi. The background of the Sámi word is not clear.<sup>38</sup> The North Russian toponyms with this formant only occur in the western periphery of the Dvina basin and in the Beloozero region (MATVEEV 2001: 188).

**-сарь** | *Кивсарь, Лансарь, Пиксарь* | meadows, islands | < 20 | < Finnic *saari* ‘island’. The meadows denoted to by this formant are situated on the islands or by the low shores of the river which form islands during the spring floods. The Finnic *saari* is without a generally accepted etymology.

**-конда** (*-канда*) | fields, pastures | approx. 20 | < Fi. *kontu* ‘house and lands surrounding it’. This word is probably a derivation of the Uralic \**konta* or \**kunta* (both forms attested) ‘group of people; administrative territory’.<sup>39</sup>

<sup>38</sup> It has been suggested that this word could be connected with an Ob-Ugrian word with a similar meaning (SAMMALLAHTI 1988: 551). This postulation is based on the assumption that in this word the first syllable \**a* is sporadically not labialised in Sámi.

<sup>39</sup> This is a new etymological version which is not to be found in standard references.



**-ванга** | *Кортеванга, Русованга, Мареванга* | approx. 10 | < Karelian *vanka* ‘meadow (on a shore of a lake or a river)’. The word is a Germanic borrowing (cf. Old Norse *vangr* ‘meadow’ < \**wanga* ‘curve’, cf. SSA III: 406; the meadows in the river valleys are typically situated at the bends of rivers).

**-ранда** | *Вочаранда, Кавкаранда, Кукранда* | approx. 10 | < Finnic *ranta* ‘shore’, a word of Germanic origin (< Proto-Germanic \**strandā*).

### 5.2. Some common bases and their origins

There are many more bases than formants in northern Russian substrate names. The bases vary much areally and there are few bases which would be present in the whole of the Arkhangelsk Region. Therefore, the list below is much less representative than the list of formants above and serves mainly as an illustration. All the examples are from the Pinega District.

As noted above, many of the etymologies for the bases are not verifiable on the basis of the characteristics of the object. Thus, the etymologisation of the bases is often more insecure than the etymologisation of the formants. However, analogical processes which affect the phonological shape of the toponym are not as common in the bases as in the formants and therefore, the bases always have their origin in the specifics of the substrate language toponyms.

The material is presented according to the probability scale presented above in section 3. Only the three most probable groups of etymologies are taken into consideration. As noted above, some elements in substrate toponyms occur both in the bases and in the formants (*-сар(ь)* ‘island’, *-ранда* ‘shore’, *-немь* ‘cape’, etc.) and these have been left aside here because they have been considered above. As there is no similar systematic presentation of toponymic bases as there is for formants (MATVEEV 2001), no figure for toponyms including a specific formant is given. One should note, however, that besides Pinega district, most of the toponymic types presented here also appear in other areas in the Arkhangelsk Region.

**A)** Toponyms belonging to toponymic types present in living languages with an etymology that can be verified by language-external facts:

**Палт-** / **Полт-** | rivers | *Полтома* two rivers, *Палтанские* fields | < Finnic \**palte* ‘slope’ (Germanic borrowing). Names denote objects characterised by hilly terrain and slopes.

**Шул-** / **Сул-** | rivers, riverside objects | *Шуланемь* cape (in two places), *Сульца* river | < Finnic \**sula* ‘melted; unfrozen’ Names denote places which remain open in the winter or open first in the spring (cf. section 6.3. on the double substitution of Finnic \**s*.)



**Хар-** | several kinds of objects | *Харанолы* field (in two places), *Харанемь* meadow | < Finnic *\*haara* ‘branch’ (Baltic borrowing); names denote geographical features which are somehow ‘branched’: one *Харанолы* is situated on a hill which has a shape similar to a horseshoe, the other is situated at a confluence.

**Юром-** | rivers | *Юрома* river (in several places) | < Finnic *\*jyr(h)ämä* ‘a deep and wide place in a river’. Names denote rivers which flow through lakes.

**Кыл(ь)м-** | brooks, rivers | *Кылма* river, *Кылмовка* spring | < Uralic *\*külmä* (> Finnish  *kylmä*) ‘cold’. Names denote objects characterised by especially cold water.

**Явр-** | brooks and rivers flowing from or through lakes | *Явроньга* ‘lake’ | < Proto-Sámi *\*jāvrē* ~ Finnic-Saami *\*jävri* ‘lake’. It is peculiar that most substrate lake names in the Arkhangelsk Region have been adapted as partial translations. Therefore, the substrate language word for ‘lake’ has been preserved only in brook and river names. They suggest that in most of the Arkhangelsk Region the word had a phonological shape close to that of Sámi *\*jāvrē* (> SaN *jävri*).<sup>40</sup>

**Торос-** | lakes, rivers | *Торосозеро* | < Sámi / Pre-Finnic *\*toras-* ‘crosswise’ (> saN *doares*, East Mari *toreš* ‘against’) Name denotes lakes which are passed through on the way to other, more important lakes.

Some names which belong to this group have etymologies not as straightforward as those mentioned above. In these cases the naming motivations are not easily understandable and, therefore, the lexemes behind the names are also not easily identifiable. In some cases investigation into place names in the living Finnic languages provides information that makes an etymological interpretation of the toponyms possible. A few cases are presented below.

**Кандело** small lake (< Finnic *\*kantelek* [> Finnish *kannel*]) ‘gusli; harp; a musical instrument’ (a Slavic borrowing). This name denotes a lake with a shape similar to a gusli. An investigation of Finnic and Karelian lake names derived from similar lexemes (NA) proves that motivation of this kind has indeed been used in naming lakes in the territory of the historical Karelian settlement.

**Варзас** a part of a river (a strait) (< Finnic *\*varkas* ‘thief’, Germanic borrowing). This name denotes a strait by the River Kuloj which forms an al-

<sup>40</sup> In fact, all the Russian substrate toponyms point to either *-vr-*, *-hr-* or *-kr-* (< *\*-kr-*) in this word (MATVEEV 2002). If these words indeed are connected to the Finnic *jävri*, the Baltic etymology for the word (< *\*jáura* ‘moor ‘moor or sea’, NUUTINEN 1989) cannot hold.



ternative and shorter waterway when moving along the river. An investigation of Finnic toponyms with similar lexemes proves that this is indeed the likely motivation for several place names derived from *varkaus* ‘theft’.<sup>41</sup> The Finnish expression *kulkea (kuin) varkain* ‘move quickly (literally: ‘like a thief’)’ is also semantically related to the motivation behind *\*varkas*-toponyms.

**Валладось** marsh < Finnic *\*valvattus* (> Finnish and Karelian *valvatus*) ‘hole in the ice that remains open’ from *valva-* ‘stay awake or open’ This name denotes an open, moist bog. Investigation of Finnic toponyms with a similar lexical content implies a common naming motivation. This word has obviously been used as a metaphor for open bogs.

**Мурд-** | *Мурдой* brook (in several places) < Finnic *\*murto(i)* ‘break (n)’, a deverbial derivation; in toponyms of Pinega district ‘whirlpool’ | The names derived from this word stem are connected to brooks which flow into the main river at narrow points where whirlpools arise. Another investigation into Finnic place names connected to a similar naming model revealed the same motivation. In Karelian there is also a dialectal word *murto* ‘whirlpool; deep water’. This clearly is a derivation from *murtaa* ‘break’. The original meaning of the word seems to have been ‘to turn back’.<sup>42</sup> One needs to be aware, however, that the word *murto* is connected to several other name types in Finnic languages as well (‘thicket; brake’; rapids’).

**B)** Toponyms which belong to toponymic types present in the living languages but which have an etymology that is not verifiable on any language-external basis

**Мамк-** | brooks, lakes, etc. | *Мамкоя* brook | < *\*matka* ‘road; passway’. In Finland, names of this kind have been given to places which were passed on the way to some important destination. As there is no information available on the traffic routes used by the pre-Slavic populations of northern Russia, it is not possible to verify whether or not a similar kind of motivation is also behind the substrate names of the Pinega basin. As this name type is common among living Finnic languages, it is likely that a similar type existed in substrate languages of the Arkhangelsk Region as well.

<sup>41</sup> Also, the name of the Finnish town *Varkaus* in Southern Savo seems to be connected to this motivation. This town is situated on an isthmus between two major lakes Kallavesi and Haukivesi near a place where big rapids Ämmänkoski flow from the previous to the latter. Travelling through the rapids by boat may have been avoided by taking a short cut through across the isthmus.

<sup>42</sup> This is also the meaning of the Mordvinian (Erzya) *murdam*s and (Moksha) *mærdem*s which have been connected with the Finnic verb with some reservations (in SSA). Also, the North Sámi *murdit* ‘retreat’, which is a borrowing from Finnish, proposes similar semantics.



**Хуð- (Xum-)** | settlements, lakes, elevations | *Хуðгора* hill, *Хумозеро* lake | < \**hiiti* (> Finnish *hiisi*, Gen. *hiiden*) ‘unholy’, originally likely ‘sanctuary; centre of a settlement’. Bases derived from \**hiiti* are typical in the present Finnic languages and they have been considered in detail in the toponymic literature (KOSKI 1967–1970). In Finland and Estonia, the place names formed from the appellative *hiis(i)* are often connected with old centres of settlements which, quite probably, had sanctuaries. The present semantics of the word seem to have developed relative to the adoption of Christianity. Also in northern Russia, some *xum-* and *xuð-*places are situated in the centres of old settlements (cf. *Хуðгора* above, section 2.3.). In other cases, this kind of correlation is not self-evident, however. It is probable that archaeological excavations could in some cases provide further support for the etymology.

**Хярз-** | brooks | *Хярза* brook (in several places) < \**härkä* ‘bull’; Toponyms formed from a word stem meaning ‘bull’ are typical of Finnic languages. However, there seems to have been a peculiar toponymic model in the substrate language of the Pinega basin: four small brooks which bare this name all have an especially strong current in spring time, while in the summer they dry up altogether. There is probably some kind of metaphoric naming motivation behind the model.

**Чухч-** | brooks; settlements | *Чухча* river (2), *Чухчамень* village < Proto-Sámi \**čukčē* (> North Sámi *čukčá*) ‘capercaillie; *tetrao urogallus*’ This toponymic etymology has been suggested in several treatises on northern Russian substrate toponymy (cf. MATVEEV 2004: 103–104). The fact that the word related to the Sámi word for capercaillie existed in the substrate languages of the territory seems well founded: the Russian dialectal *чухарь* and the Komi dialectal *čukči* which both mean ‘capercaillie’ have, most likely, been borrowed from substrate languages of the territory.<sup>43</sup> However, the naming motivation for the *чухч-*places can hardly be verified in most cases. Moreover, there are other problems related to the interpretation of Sámi elements in substrate names (see section 6.1.).

**Нюхч-** | rivers, settlements | < Proto-Sámi \**hukčē* ‘swan’ (> North Sámi *njukča*) As with place names formed from \**čukčē* it is not possible to verify or falsify this old toponymic etymology (originally suggested by CASTRÉN, cf. MATVEEV 2004: 94–95) on the basis of language-external facts.<sup>44</sup>

<sup>43</sup> The development of the Russian word has certainly been affected by *злухарь*, the literary Russian designation for capercaillie.

<sup>44</sup> It has been suggested that this kind of bird names may also have been used as a sort of totem names (MATVEEV 1986). At the present stage of the research, this hypothesis is quite speculative but may well prove to right in principle.



C) Toponyms formed from identifiable Uralic lexemes not used in toponymic formation in living languages (or used only according to some other naming motivation)

**Кыч(ас)-** | several kinds of objects | *Кыча* lake *Кычас* lake, *Кычверетия* a passway between marshes < \**kicca(s)* ‘narrow’; the objects denoted to are characterised by their narrowness. Living Finnic languages lack a similar naming model.

**Ухм- (Охм-)** | rivers, lakes, objects related to bodies of water | *Охтома* river (< \**ukti* ‘way; passway’ (> Khanty V *оҕат* ‘track’, etc., Mansi KU *āχt* id., etc., Hungarian *út* ‘way; road’). As noted by MULLONEN (2002: 208–217) toponyms with this base denote rivers or water routes which have a narrow passway by land to other water systems (Ru. *волоок*). It is probable that in these toponyms a word present in the Ugric languages and meaning ‘passway’ or ‘road’ has been preserved (SAARIKIVI 2004c: 349). This word has no cognate in present Finnic or Sámi but it seems to have existed in the extinct languages of the Finnic and Sámi type spoken in northern Russia.

### 5.3. Old Finnic personal names and the northern Russian substrate toponymy

So far, the northern Russian substrate toponyms have been studied almost exclusively on the basis of appellative lexicon. However, the present Finnic languages also have a substantial number of toponyms formed from personal names. These are especially characteristic of settlement and field names. In Finnic languages, toponyms derived from personal names constitute approx. 10% of the total number of toponyms (KIVINIEMI 1990: 143–145). In settlement names their number may be as high as 50% (MULLONEN 1994: 85–86).

In northern Russia, only some isolated examples of substrate toponyms derived from personal names have been presented in the toponymic literature so far (see SAARIKIVI 2003). This is partly due to a lack of historical documentation. There are few documents which name individual pre-Slavic settlers in northern Russia, and probably not a single document that would with certainty connect a particular individual to a specific place. Further, the system of old Finnic personal names has been described fairly superficially.

Only a limited number of Finnic pre-Christian personal names has been preserved in historical sources. It is clear, however, that in a similar manner to toponyms, many Old Finnic personal names have consisted of two parts (*Kauko/valta*, *Iha/lempi*, *Vihta/mieli*)<sup>45</sup> or have been based on participles



<sup>45</sup> *Kauka-* is modern Finnish for ‘lengthy, long’, *valta* ‘power; might’, *lempi* ‘love’ and *mieli* ‘will; desire’. *Iha* and *vihta* are nonexistent in modern Finnish. The

(*Valittu* ‘choiced’, *Lemmitti* ‘beloved’, *Toivottu* ‘hoped’). Quite likely, the first part of the two-part names was also used on its own. It may be assumed that when toponyms were formed from personal names, the generic of the name was eliminated and the first part of the name began to be used as the specific of a derived parallel toponym (*Ihamieli* ‘personal name’ + *mäki* ‘hill’ > *Ihamäki*, *Kaukovalta* ‘personal name’ + *la* ‘locative suffix’ > *Kaukola* settlement).

In living Finnic languages toponyms derived from personal names are most typical in settlement names, quite typical in names connected with agriculture and quite atypical although not nonexistent in hydronyms. The probability of an etymology based on a personal name also follows this form. However, because of a lack of literary sources, all the toponymic etymologies based on personal names would belong at the maximum to group 2 on the probability scale. There is, to be precise, nothing in the places themselves that could verify or falsify an etymology based on a personal name.

In some cases, it is hard or even impossible to decide whether a substrate toponym was based on a personal name or a corresponding appellative. Thus, it is not clear whether toponyms with the bases derivable from the Proto-Finnic *\*repoi* ‘fox’ (e.g. Pinega settlement name *Revomurga*,<sup>46</sup> cf. MATVEEV 2004: 63) can be connected to the appellative meaning of the word or to the personal name based on the appellative and attested in literary sources (cf. STOECKE 1964: 64).

In the following, some northern Russian place name types have been etymologised on the basis of Finnic personal names (some of them were presented earlier in SAARIKIVI 2003).

**Ихал(ь)-** | settlements, meadows, brooks, etc. | *Ихальнемь* meadow, *Ихала* river, *Ихалово* village, etc. (see MATVEEV 2004: 37–38) | < personal name *\*Ihala*. The one-time existence of this name in northern Russia is verified by the Novgorod birch bark letter 249, which includes the personal name *Игала* (ZALIZNYAK 2004: 623–624).<sup>47</sup> Although the Finnish dialectal and Karelian adjective *ihala* ‘lovely; delightful’ also exists, it is probable that most of the northern Russian substrate names with this base are derived from personal names. There are many personal names derived from *iha* ‘delight’ (*Iha-*

---

former has, however, survived as a derivation *ihana* ‘lovely; delightful’. *Vihta* is a name element with a likely Germanic origin.

<sup>46</sup> *Мурга* is a Russian dialectal geographical appellative meaning ‘pit caused by erosion’ (cf. SAARIKIVI 2004a: 196–197).

<sup>47</sup> This name has already been identified as Finnic already by HELIMSKI (1986).



*lempi, Ihamieli, Ihamuoti*),<sup>48</sup> and *\*Ihala* certainly also belongs here. The same name is also preserved in the Finnish surname *Ihalainen* (SNK 148, cf. also SAARIKIVI 2003: 144).

**Кавка-** | *Кавкола* village (in the mouth of Dvina) < personal name *\*Kaukoi*. A similar name element has been used as a first component of several pre-Christian Finnic personal names (*Kaukomieli, Kaukovalta, \*Kaukohalu*,<sup>49</sup> etc.) and it has also been preserved in several Finnish surnames (*Kaukinen, Kauko, Kaukonen*, etc., SNK 207–208).<sup>50</sup> Some substrate names with the lexeme *kauka-* can be connected with the appellative semantics of the element *\*kauka* ‘distant, remote’, originally ‘long’ (cf. MATVEEV 2004: 38).<sup>51</sup>

**Ракул-** | settlements, bodies of water | *Ракула* settlement, *Ракулка* river | < personal name *?\*Rakkoi(la)*. This frequent northern Russian settlement name type has been interpreted as Finnic though without a true etymology by MATVEEV (1999: 86). It seems likely that it was based on the Karelian personal name *\*Rakko(i)* which has been preserved in some literary sources and in Finnish surnames *Rakkola* and *Rakkolainen* (SNK 521).

**Вухт-** | village, branch of a river | *Вухтово* (< *Вухтуу*, a form attested in early documents) village, *Вухтовский* river branch | < personal name *\*Vihto(i)*. The village name *Vihtovo* in the Pinega District is one of the oldest in the Dvina basin, attested even in 1137. It is, most likely, connected with an element attested in several old Finnic personal names (*Vihtimeeli, Vihtari, Vihtiä*, STOEБKE 1964: 105–106). Also, this personal name has been preserved in the birch bark letter 2 (anthroponym *Вухтимасъ*) and in the Finnish surname *Vihtonen* (SNK 744).

**Хим-** | meadows, bodies of water | *Хима* river, *Himasora* brook, etc. | < personal name *\*Himo(i)*. The same name element occurs in compound personal names *Himopää, Himatoinen, \*Himottu*, etc. which have been preserved in old literary sources (STOEБKE 1964: 20–21). Likely, the personal name *Гьмуй*, mentioned in birch bark letter 403 also belongs here. The same element has also been preserved in the Finnish surname *Himanen* (SNK 120).

<sup>48</sup> Literally ‘lovely form’. The is name has probably meant, approximately, ‘good-looking’.

<sup>49</sup> This kind of previously nonattested personal name most likely appears in the Novgorod birch bark letter 249 (*У Кавкагала*), referred to above.

<sup>50</sup> Even today there is a christian name *Kauko* in Finland, although this is a formation of the period of national romanticism.

<sup>51</sup> MATVEEV (ibid.) has also connected the name of the village *Кавкола* to the Finnish *kaukalo* ‘vat’ and its Finnic cognates. This is extremely unlikely, because no similar toponyms are attested in the present-day Finnic area.

The appellative *himo* means ‘lust; desire’ and it is likely that this meaning is also behind the personal names.

**Айн-** | *Айново* village in the Pinega District | < personal name \**Aino(i)* (not attested in literary documents). The literary meaning of the name was probably ‘sole; the only one’ (fi. *ainoa* ‘the only one’). Further, such names as these have been preserved in Finnish surnames (*Ainas*, *Ainalinen*, *Ainoinen*).

The examples above demonstrate that Finnic personal names are useful in the search for etymological cognates to northern Russian substrate names. While it has been considered an out-dated tradition in Finnish toponymistics to explain unintelligible place names by loosely suggesting that they may include old personal names, of explanations of this kind should not be categorically rejected. They can be proposed by stricter criteria than those suggested by previous scholars. Especially in cases in which a common element occurs both in surnames and several individual place names connected to settlements, does the reconstruction of an old personal name seem possible. Many Finnic personal names have also been preserved in the Novgorod birch bark letters and this substantially enhances the credibility of some of the comparisons above. In addition, old Finnic personal names have been preserved in surnames and toponyms which denote settlements and belong to types typically derived from personal names (most notably, toponyms with word final *-ла*, a formant that originates in Finnic settlement name suffix and *-ev(o)/-ov(o)*, Russian settlement name suffix)].

In addition to old Finnic personal names, it also seems likely that personal Christian names have survived in the substrate toponyms of the Dvina basin (*Лукомень* < ?*Лукий*, *Иванемь* < ?*Иван*, *Юрола*, *Юрьемь* < ?*Юрий*, etc.). There would be nothing strange in 14–16th century Finnic settlers in the Dvina basin adopting the Christian name system. Similar anthroponyms and toponyms are today commonplace among the Finnic people of northern Russia.

#### **5.4. Appellative substrate vocabulary and substrate toponyms**

Many words present in substrate toponyms also occur as appellative borrowings. The borrowing of toponyms and geographical vocabulary are related phenomena which both typically occur in the case of language shift (see in detail SAARIKIVI 2000; AIKIO 2004). Place names and geographical appellatives are learned in a similar manner, while learning the concrete objects they denote.

As noted already by generations of scholars, most of the appellative borrowings in northern Russian dialects are of Finnic origin. In addition, there are a few borrowings considered to be Sámi, some Komi and Nenets borrowings



and vocabulary from unidentifiable but, most probably, Uralic sources. Among the frequent semantic fields of Uralic borrowings are words related to geography, weather conditions and northern means of livelihood such as fishing, hunting and reindeer herding (MYZNIKOV 2004: 78–248).

There are two groups of appellative vocabulary that can be considered linguistic substrate in the sense that they have belonged to the vocabulary of an extinct language in a specific area. These are 1) vocabulary that besides appellative use also appears in substrate toponyms and 2) vocabulary that denotes strictly local concepts and has a narrow distribution in dialects. For example, the well-known Finnic borrowing *лахта* ‘bay’; also (through metonymy): ‘marsh; moist place; meadow’ (< \**lahti* ‘bay’, KALIMA 1919: 151)<sup>52</sup> has a wide distribution in North Russian. In the Pinega District, it forms many Russian toponyms that consist of an adjective attribute and a geographical appellative (*Великая лахта* ‘large bay’, *Грязная лахта* ‘soiled bay’, etc.). As it also occurs as a formant in substrate toponyms (*Куклохта* meadow, *Киглохта* village, *Ролахты* bay) we know that it has belonged to the extinct Finnic vernacular of the Pinega basin and has not spread there through other Russian dialects. Similar terms with a wide distribution in Russian dialects, but which are fixed in the substrate toponyms of the Pinega District are *луда* ‘rocky islet’ (< Finnic *luoto* id.), *каска* ‘young woods’ (< Fi. *kaski* ‘burnt-over clearing; woods that grow in it’), *виска* ‘brook that flows out of a lake’ (< ?Fi. *vieska* ‘current in rapids’)<sup>53</sup>, *щелья* ‘hill or steep bank by a river’ (< \**selkä* ‘ridge (originally: ‘back’)’<sup>54</sup>, etc.

The other group of geographical terms of substrate origin has a very limited distribution in dialects. Typically, these are words which denote the geographical features of some specific microterritory. They may denote only to a few places and, therefore, are used in a manner close to the use of toponyms. Thus, the dialect word *мурга* ‘funnel-like pit caused by erosion’ is only attested in the Pinega dialect of Russian (SRNG 18: 353) and the adjacent Udora dialect of Komi (KESKJ 179). This is natural in that the objects it denotes are uncommon in most of northern Europe. In the Pinega re-

<sup>52</sup> The Russian word could also have originated from the pre-Finnic \**lakti*.

<sup>53</sup> This etymology (proposed by the author of this article in SAARIKIVI 2004a: 196) is insecure because the Finnish dialectal *vieska* has a narrow western distribution and the meanings of the Russian and Finnic words are different. According to another also problematic version, this word is a Komi borrowing (REW I: 204; KESKJ 58).

<sup>54</sup> The initial *ш* which occurs only in some dialects (the other dialects have *ш*), is probably the result of folk etymology. The word was contaminated with the Russian *щель* ‘gap; hole’ (rivers with steep banks flow through gorges, see SAARIKIVI 2004a: 197).

gion, this word is connected to pits caused by the rapid erosion of soil consisting of karsts. The fact that the word belonged to the substrate language of Pinega basin is reinforced in that *мурга* also occurs as a formant in at least one substrate toponym (*Ревомурга*, a settlement name, cf. above).

Another group of words which seems to originate in the substrate language is used in toponyms not as formants or bases of substrate toponyms, but only quite alone (as the only lexeme in toponym) or in conjunction with the Russian adjective attribute. In these cases, the dialectal distribution and the phonological shape are the main criteria in classifying the words as local substrate borrowings. Thus, Russian dialectal *койдома* ‘passable marshland’ has been attested only in Pinega and some nearby districts. The word seems to be connected with Finnic *keidas* ‘high place on a swamp, etc.’ which, in turn, is a Germanic borrowing (< *\*skaiða-z* ‘passage, distance, interval’, SSA). There is no word that would directly correspond to the Russian dialectal *койдома* (< likely *\*kaitama*) in the Finnic languages, but as we know that the Finnish *keidas* had *\*ai* in the first syllable and geographical terms with a derivational suffix *-mA* (or *-mO*) are commonplace in Finnic languages (Hakulinen 1979: 130–131), it is quite possible that in the extinct Finnic dialect of Pinega, a word *\*kaitama* ‘passable swamp’ has existed (SAARIKIVI 2004a: 195–196).

A similar case, although with a somewhat wider dialectal distribution is the *мег* ‘bench of a river’ which could have been borrowed from *\*mäki* ‘hill’ (see VESKE 1890: 164). In modern Finnic, *mäki* only means ‘hill’ but in the Finnic substrate language of the region, the semantic shift ‘hill’ > ‘bend of a river; promontory’ would appear to have taken place. This shift would be explicable in that *mäki* would have first developed the meaning ‘a high place by a river’. A similar semantic shift has occurred also in Slavic: the cognate of the Russian *берег* ‘shore’ (< PIE *\*bhergh-*) means ‘hill’ in Germanic (cf. German *Berg*).<sup>55</sup> The presumed semantic shift can be further supported by the use of the word in the Pinega dialect. It is frequent in expressions such as *у̀дму через мег* ‘walk through a bench of a river (i.e. not by the coastline but over land)’ and *на мегу* ‘at the bend in a river (i.e. not by the shoreline)’. Moreover, the Finnic *mäki* ‘hill’ is probably nonexistent (or very rare) in Finnic substrate toponyms of the Dvina basin although it does belong to most common generics in all of the Finnic languages. As most of the other common generics of Finnic are otherwise present in Dvina basin place

<sup>55</sup> The word *мег* has also been borrowed into Komi dialects, probably from the substrate languages of the Dvina basin. The etymological explanation given by KESKJ (~ *ud mog*, saN *mohkki*, p. 171) is rejectable on phonological grounds (the vowel correspondences are not regular).



names, the absence of *mäki* would be surprising, especially if one takes into account that it is among the most common geographical appellatives in the toponym formation of many Finnic languages (cf. KIVINIEMI 1990; MULLONEN 1994: 26).

Thus, there are borrowings in North Russian dialects, which have probably originated in extinct Finnic languages with no exact parallels among present-day Finnic idioms. As many of them denote geographical concepts and are used in toponym formation, the study of appellative substrate vocabulary is intimately connected with the study of substrate toponymy. One should note, however, that those toponyms including only a geographical appellative should be classified as Russian and not substrate toponyms.

## 6. Ethnical interpretation of northern Russian substrate toponyms

### 6.1. The dating of Russian colonisation in the Dvina basin

The substrate toponyms of the Dvina basin reveal no traces of such Slavic sound shifts as *polnoglasie*, elimination of nasal vowels or disappearance of the yers. This clearly points to the fact that Slavic spread to this area later than it did to the vicinity of the Gulf of Finland where these phonological phenomena are present in some toponyms. It is not clear, however, from the substrate toponymy where the even approximate borders of these sound shifts are to be found. Many scholars have pointed to such Novgorod Region toponyms as *Мста* (< \**Mustajoki* and *Нарова* < *Narva*, cf. AGEEVA 1989: 220–221) which presumably represent reflexes of these sound shifts. Moreover, many Novgorod Region river names of probable substrate origin seem to end in a consonant (*ibidem.*) whereas river names of this kind in the Arkhangelsk Region are rare. This suggests that Novgorod Region names ending in a consonant have had word final yers. One should note, however, that the main bulk of appellative Finnic borrowings in Novgorod dialects are more recent (MYZNIKOV 2004: 261–263) and this leads to the conclusion that Finnic-Slavic contacts in this area lasted for a long period. Also MULLONEN (2002: 43–51) has pointed to some toponyms from the Svir' basin which seem to be have been borrowed before the disappearance of the nasal vowels and yers (*Винницы* < *Veps Vingl*, *Свирь* < \**Syväri*).

The disappearance of the yers has been dated at 1150–1300 by ZALIZNYAK (2004: 59–62). As there are no traces of yers in the toponyms of the Dvina basin, one has to admit that the entire Dvina basin must have been linguistically overwhelmingly Uralic until the beginning of the 14th century.

The distribution of different morphological adaptation types of substrate toponyms is probably connected to the different russification patterns of Uralic populations. It has been demonstrated that the partial translation pat-

tern (cf. section 3.2. above) has spread into those areas in which the Slavic population came from Novgorod (GUSELNIKOVA 1994: 12). MULLONEN (2002: 128–132) has convincingly demonstrated that the distribution of brook names with the formant *-ou* and Russian partial translations with the ending *-пучей* ‘brook’ correlate with the Ladoga-Tikhvin and Onezhskaja group of Russian dialects and the archaeologically defined border of the early (prior to 1000 AD) and late (after approximately 1250 AD) Slavic colonisation of the Svir’ basin. She suggests that the full adaptation of toponyms would have been connected with the Slavic migration to the Svir’ basin, while partial translations would be the result of a slow russification of the indigenous Uralic population through language shift.

It is not clear yet whether similar correlation patterns can be observed elsewhere, also. One should note, however, that correlations of this kind are not universal. For instance, in the Finnic-Sámi contact zone (inner Finland) all substrate toponyms are adapted as partial translations (cf. ANTE AIKIO’s article in this volume).

### **6.2. Identification of substrate languages: were there Sámi in the Dvina basin?**

Most of the examples referred to above are from Finnic languages. However, all scholars agree that many toponymic types of northern Russia cannot possibly be explained solely on the basis of the Finnic languages. It has been continuously proposed since Castrén that besides Finnic tribes, also the Sámi inhabited northern Russia. As noted above, this argument was based on toponyms which include lexemes present in Sámi languages. It finds limited support in ethnotoponyms and there are also few fragments of oral tradition which could be related to the Sámi (see MATVEEV 2004: 192–193 and article by A. K. MATVEEV in this volume).

However, the northern Russian place names indicate very peculiar kinds of “Sámi” languages. Those Sámi languages known to present linguistics have a large amount of vocabulary without Uralic cognates or loan etymologies (cf. ITKONEN 1948: 16–26). These vocabulary layers can be considered borrowings from from extinct Paleao-European substrate languages (for details see AIKIO 2004, SAARIKIVI 2004a). The frequent but unetymologisable Sámi geographical terms (North Sámi forms given) *njárga* ‘cape’ (< *\*ńarke*) and *geađgi* ‘stone’ (< *\*kēðkē*) occur in toponyms only to the west of the Dvina basin, and the area of distribution of some other central terms (such as *bákti* ‘rock’ [< *\*pāktē*], *roavvi* ‘place where there has been a forest fire’ [< *\*rōvē*], *vuotna* ‘fjord’ [< *\*vuongē*], etc.) is even more northern and western (SAARIKIVI 2004b: 206–210). Thus, important layers of vocabulary present



in Proto-Sámi and its offsprings are nonexistent in the “Sámi” place names of the Dvina basin.

Further, toponyms with phonological and morphological developments characteristic of Sámi languages do probably not exist in most of the Arkhangelsk Region. Thus, the attribute form of the adjective *guhkki* ‘long’, *guhkes* (< Proto-Sámi *\*kukēs*) which occurs in several Sámi substrate origin lake names in Finland and Karelia, is nonexistent in the substrate toponyms of the Dvina basin (SAARIKIVI 2004b: 202). This is symptomatic, because the existence of a separate attribute form of an adjective is a characteristic and innovative feature of the Sámi languages. The fieldwork by the author also implies the conclusion that, in the Pinega basin, toponyms with the base *kuk-* characterised as Sámi by MATVEEV (2004: 185), are more likely connected to the Finnic *\*kukku(la)* ‘hummock’.

The traces of regular Sámi sound shifts have in many cases been flushed away by the Russian adaptation of the place names (cf. results of the Sámi vowel shifts *\*i, \*e, \*ü > (North Sámi) a, \*a > (North Sámi) uo*, etc., and the substrate language—Russian sound correspondences *\*a, o ~ o, e, a ~ a*, etc. referred to above). However, some Proto-Sámi vowel shifts are attested in toponyms in the western parts of the Arkhangelsk Region (op.cit 196–198, cf. toponymic types *лумб-* ‘small lake’ and *еле-* ‘upper’). There are also examples of the Sámi development *\*ś > ć* in some appellatives (cf. Russian dialectal appellative *чильма* ‘an open place in a marsh’ (< *\*śilmä* ‘eye’, MATVEEV 1978)<sup>56</sup> and toponyms with the base *чолм-* ‘strait’ (< *śolma*, MATVEEV 2004: 316; SAARIKIVI 2004b: 197–199).<sup>57</sup>

The picture of the substrate languages in the Dvina basin becomes even fuzzier if one takes into account that elements characterised as Sámi by generations of scholars, combine with elements which may only be characterised as Finnic. This results in toponyms which are certainly Uralic, but which are difficult to interpret from the point of view of Uralic linguistic taxonomy. Thus the specific of the name *Чухчемена* has been interpreted on the basis of the Sámi *\*ćukcē* (> North Sámi *ćukcá*), ‘capercaillie’ whereas the generic of the name is without doubt connected to the Finnic *\*neemi*

<sup>56</sup> This word is connected to Finnic lexical convention (Finnish form given) *suon-silmä* literally ‘marsh-eye’ = ‘an open place in the marsh’ from *silmä* (< *\*śilmä* ‘eye’). This convention is nonexistent in Sámi languages, while the offspring of PU *\*śilmä* (> saN *čálbmi* ‘eye’) is otherwise present. The word also lacks the Sámi vowel developments.

<sup>57</sup> Note, that in the latter article it has been argued that this word may also be offspring of Pre-Finnic *\*ćolma*. The implications of northern Russian toponyms for the history of Finnic and Sámi affricates are discussed below in 6.4.

‘promontory; cape’, which, in turn, is nonexistent in Sámi (MATVEEV 2004: 225–226, cf. also names like *Нюхчалакша, Шубоя, Шубматка*, etc.). It seems likely that names of this kind are not Sámi-Finnic partial translations either, because no Finnic language has the sound combination *-hč-* (with the exception of some late Votic cases). Thus it seems justified to suggest that we are dealing with toponyms from extinct languages which shared lexical features of present Finnic and Sámi branches of Uralic languages (see, however, A. K. MATVEEVs differing opinion in his article published in this volume and MATVEEV op.cit.).<sup>58</sup>

Moreover, as noted above, there are also northern Russian toponymic types etymologisable on the basis of Uralic languages which are, at least apparently, neither Sámi nor Finnic. For example, place names with the bases *yxm-* and *кыч-* or the formants *-capa* or *-nana* are certainly Uralic, but they cannot be labeled according to the present Uralic branches. This also implies that the toponymic types referred to by MATVEEV with close resemblances in the Sámi languages (cf. *нюхч-, чухч-, торос-* above; see MATVEEV 2004: 210–231 for more types) did not necessarily originate in a language which should be characterised as Sámi in the present sense of the word.<sup>59</sup> Moreover, many of MATVEEV’s etymologies are uncertain (they belong to categories 2, 3 and 4 on the probability scale presented above) and some could well be interpreted as Finnic (cf. toponymic bases *налд-* < *\*palt(t)e-* ‘slope’ [and not (North) Sámi *bealdu* ‘field’, MATVEEV 2004: 95], *чыга* [< ??Vepsian *čuga* ‘corner; spot’ or Vepsian *\*čuhu* ‘hill’, a lexeme reconstructed on the basis of toponymy, MULLONEN 1994: 56–57] and not Sámi *\*čokke* ‘top of the hill’, cf. *ibid.* 102–103], *кыч-* (< Finnic *\*kukku(la)* and not Sámi *\*kukkē* ‘long’, cf. *ibid.* ).

Instead of speaking of Sámi toponyms in the eastern and central Dvina basin, one should probably speak of toponyms which share some phonological and lexical features with the Sámi languages. They seem to have originated in Uralic language forms which also underwent the sound shift *\*š > č* and had several lexemes in common with the Sámi languages. However, proba-

<sup>58</sup> Sámi *čukcá* is without Uralic cognates. This word presents a phonotactic structure that has no regular correspondence in present-day Finnic (first syllable *u* + middle consonant cluster, second syllable *á*). Therefore, it is likely that even in Sámi, this word is a Palaeo-European substrate borrowing. Komi *čukč̄i*, referred to as a cognate word in UEW and KESKJ is probably a borrowing from substrate languages of the Dvina basin.

<sup>59</sup> An especially peculiar case is the base *нюхч-* which probably is connected to a word meaning ‘swan’ that is present in many Uralic branches. Words belonging to this connection have many irregular sound correspondences (Sámi has irregular word initial shift *j > i*).



bly not one of the central geographical appellatives which today differentiate Sámi toponymic systems from Finnic systems was present in these languages. The hypothesis that there were substrate languages of non-Finnic and non-Sámi character is further supported by the fact that the historical sources mention several tribes without parallels among the present Uralic peoples.

In the western parts of the Arkhangelsk Region, there seem to have been substrate languages closer to modern Sámi in some respects—two good candidates for areas with such a substrate language are the Beloozero region and the Lower Onega region (see MATVEEV 2004: 114–131; 181–186). But even these languages were lexically not similar to modern Sámi. Place names in the Dvina basin point to a dialect continuum in which lexemes and innovations present in the modern Sámi languages increase to the west and diminish to the east. Where exactly the substrate toponymy should be labeled as Sámi is a question that cannot be unambiguously answered.

At present the question of non-Finnic substrate languages in the Dvina basin is far from settled. Further, the hypothesis that there were Sámi in the Dvina basin may find support when the etymological study of place names in the area proceeds. Most likely, this must be solved by areal investigation of toponyms. It is sure, however, that possible Sámi languages in this area were linguistically much less similar to the modern Sámi languages than Finnic tribes in the area were to modern Finnic.

### 6.3. Identification of Finnic tribes

In research history, the Finnic tribes of the Dvina basin were considered Karelian (CASTRÉN 1844, KIRKINEN 1963), Veps (HAAVIO 1965, PIMENOV 1965) and lately Karelian, Veps and other Finnic (MATVEEV 2004: 194–204). In ethnic interpretation of place name material, ethnotoponyms have dominated: the *чудь* have mainly been interpreted as Veps, while the idea that there were Karelians in the Dvina basin was based on ethnotoponyms derived from the ethonym *Корела*.

In addition to Russian ethnotoponyms, the most promising methods in identifying the Finnic substrate languages are a search for vocabulary present in some Finnic languages and nonexistent in others, and a search of naming models historically productive in specific Finnic languages and nonexistent in others. The third method available in differentiating Sámi toponyms from Finnic ones, a search for traces of regular phonological shifts, is not easily applicable in the case of Finnic toponyms, because only minor sound shifts differentiate individual Finnic languages and even their traces have often disappeared, especially if the toponyms have been borrowed from one



Finnic language into another. However, some toponyms still hint at substrate languages with specific phonological characteristics.

As in the case of toponyms characterised as Sámi by generations of scholars, the distribution of lexemes, naming models and phonological shifts characteristic of individual Finnic languages is not easily interpretable in ethnic terms. Thus, in the lower Pinega basin where there are *корела*-ethnotoponyms, no definite traces of the most frequent Karelian toponymic term *lampi* ‘small lake’ are attestable. This state of affairs may, of course, be connected with the small number of lakes in this area, but also frequent Karelian name models such as *karsikko* ‘memorial tree’, *ryhjä* ‘centre of a village’, *nilos* ‘smooth; slippery’, *haiseva* ‘stinking’ (concerning these models see KUZ’MIN 2004, VAHTOLA 1980), etc., are nonexistent in the area. This signifies substantial differences between the languages of Karelians in inner Finland and present-day Karelia, and the probable Karelians in the Dvina basin.

Some name types traditionally characterised as Karelian are present in the Pinega basin, however: *cepz-* (< *\*särki* ‘roach’), *лан-* (< *?\*lappi* ‘Sámi; North Karelian’) and probably even *кватъ-* (< *kuadjad* < *\*kaatiot* [~ Russian dialectal *zamu* ‘pants’]). The last one of these also points to a Karelian sound shift *aa* > *ua* in first syllable.<sup>60</sup> Another possible Karelian phonological shift present in Pinega toponymy is *s* > *š*, which seems to occur in the base *уул-* ‘unfrozen’ (< Karelian *šula* < Proto-Finnic *\*sula*).

In the same area, many substrate toponyms have a phonological shape close to Veps. Thus, the bases *варгас* and *ламбас* (see above 5.2.) have preserved the word internal consonantism of Proto-Finnic which in other Finnic languages has changed as a result of consonant gradation (*\*varkas* > Finnish *varas* [Gen. *varkaan*], *\*lampas* > Finnish *lammas* [Gen. *lampaan*]). At the same time, in the substrate names there are no traces of voiced stops, a phonological feature characteristic of Veps.<sup>61</sup> Also a couple of lexemes nonexis-

<sup>60</sup> KIVINIEMI (1977: 200) has identified this as a metaphoric Karelian name type used to refer to lakes which consist of two branches or two lengthy bays. *Кватъозеро* indeed has this kind of a form. If this really is a Karelian name, it has to be supposed that diphthongisation of the *aa* had happened in the substrate language. This is a development characteristic for Karelian only of all the Finnic languages. For the reasons discussed above, identifying *Кватъозеро* as a genuine Karelian place name is, however, premature.

<sup>61</sup> The voiced stops in the toponyms *Ламбас*, *Варгас* etc. are a result of the phonological adaptation of substrate names in to Russian (cf. section 4.1.). Voicing of stops is probably a relatively new sound shift in Vepsian. It is not attested in the place name material of *Писцовая книга Обонежской пятинны* from the end of the 15th century.

tent in Karelian, but present in Veps appear in Pinega toponyms: *Чуга* (< Veps *čuga* ‘angle; spot’ or *\*čuhu*, *\*čuhak* ‘hill’, cf. MULLONEN 1994: 56–57), *Пурдева* (< Veps *purde* ‘spring’). However, these combine with words which are nonexistent in the living Veps toponymy (such as *\*hattara* ‘cloudlet’, in dialects: ‘bush’ > *Хатара*, *Хатармень*, *\*laama* ‘wide place at a riverrun’ > *Ламозеро*, *\*hetteh* > Finnish *hete*] ‘spring’ > *Хетельга*).

Thus, while the overall selection of lexemes and the phonological characteristics of the substrate toponyms in the Pinega basin are probably closer to modern Veps than to modern Karelian, the substrate toponyms of the region cannot easily be labeled either Veps or Karelian. In addition, some toponymic types such as settlement names with the formants *-нала* and *-вера* have their closest parallels in the southern group of the Finnic languages. The fact that the vowel combination *e — a* has been substituted uniformly in Russian substrate toponyms and yielded the central vowel in Southern Finnic is also a remarkable parallel with Southern Finnic and the substrate languages of the Dvina basin. Furthermore, some northern Russian toponyms also suggest a substrate language that would have preserved the diphthong *\*ai* in cases where most of the Finnic languages have secondary *ei*, cf. the appellative *койдома* (section 5.4.), and toponyms with the base *хайн-* (< ??*\*haina* ‘hay’, a Baltic borrowing [> Finnish *heinä*], see MATVEEV 2004: 73–74). A similar retention occurs in South Estonian and Livonian.<sup>62</sup>

Thus, there are features of various Finnic languages in the substrate toponymy of the Dvina basin. In addition, some words present in northern Russian toponymy can be identified as Finnic, but they appear anomalous from the point of view of closer identification of the substrate language. Thus, the formant *-нала* ‘village’ has no appellative cognate anywhere in living Finnic and the frequent formant *-сара* ‘brook’ can only be compared to a marginal Finnish and Karelian dialect word which is not common in toponyms in any living language. Also, toponymic types such as *кыч-* ‘narrow’ and many geographical appellatives (*курья* ‘lengthy bay’, *рада* ‘marsh that grows low woods’, *койдома* ‘passable marshland, *мег* ‘bend of the river’, cf. section 5.4. above) do not point to any living Finnic language but rather, to a Finnic idiom lexically different from all present-day Finnic languages.

Some facts suggest that the Finnic population of the Dvina consisted of several different linguistically definable groups. Thus, in the Pinega district

<sup>62</sup> The etymology *хайн-* < *\*heinä* (MATVEEV 2004: 73–74) is not the most reliable. In the Pinega district, there are four names with this base and none of them is connected to a place in which hay now grows.

there are two parallel toponymic bases *сул-* and *уул-* with a similar motivation (< *\*sula* ‘melted; unfrozen’). This suggests that the Finnic population probably arrived in the territory in several waves, in a similar manner to present-day Finland where competing toponymic patterns of different Finnic tribes often exist side by side in the same region (cf. VAHTOLA 1980; KIVINIEMI 1971).

As there are historical sources suggesting a Karelian presence in the Dvina basin in the 15–16th centuries (cf. KIRKINEN 1963), it seems reasonable to assume that some relatively modern Karelian toponyms of the Dvina basin bear witness to Karelian settlers who arrived in the territory just before or simultaneously with the Slavic migrants from the southern Novgorod lands, (probably at a time when the Karelian sound shift *aa* > *ua* had already occurred). This is in accordance with the views presented by MATVEEV (2004: 198–201) that Karelians settled along the lower reaches of the river valleys, whereas the Veps diffused into the forests at the southern edge of the Arkhangelsk Region.<sup>63</sup> However, this line of reasoning does not answer the question as to why several frequent Karelian toponymic types did not spread into the Dvina basin, or at least, not into the Pinega district. Perhaps this is related to the late appearance of Karelian settlement (probably at a time when several toponymic types present in Karelian toponymy had lost their productivity). It may also be partly due to the geographical differences between Fennoscandia and the Dvina basin.

Before these late Finnic newcomers, tribes speaking an archaic Finnic language forms with the diphthong *\*ai* instead of *ei* in first syllable, lack of consonant gradation and likely also a mid-central vowel similar to the Estonian *õ* in the phoneme inventory lived in the Dvina basin. It is not clear, how uniform these Finnic language forms were. The fact that there are numerous tribe names attested in the historical literature suggests that there may have been many Finnic tribes without a common ethnonym and identity. The speakers of these Finnic languages employed some toponymic types with no close parallels in the present Finnic languages. However, some of them probably used the same ethnonym (*уѣдѣ*) of themselves as some groups of Veps in the 19th century.

---

<sup>63</sup> MATVEEV (op.cit.) also refers to the fact that in the Beloozero region, there is at least one clear Vepsian sound shift which occurs in the toponymy, namely, *is* > *iš*. If this is correct, it would well correspond with historical sources pointing to a Veps settlement in Beloozero (e.g. Russian primary chronicle).



#### 6.4. Permian and still other layers of substrate toponyms

The Permian traces in the toponymy of the Dvina basin are somewhat minor and have therefore, been left mainly untreated above. There are some areas with a substantial number of Permian substrate names such as the lower Vycheгда, which was likely inhabited by Permian tribes in the Middle Ages (TURKIN 1971). It has also been proposed that the *тоймичи погане* mentioned several times in the Chronicles could have been a Permian tribe.<sup>64</sup> In the Pinega basin Permian toponyms, though quite common, seem to form a more recent superstratum layer on Finnic and other layers of substrate toponymy. This is in accordance with a hagiographical account *Житие Стефана Пермского* which mentions Komis who refused to convert to Christianity and moved from Vycheгда to Pinega in 14th century. Most certainly, the Pinega basin has been one of the key areas of late Finnic-Permian language contact as there are many Finnic borrowings in the neighbouring Udora dialect of the Komi language (LYTKIN 1967).

The Permian-Finnic linguistic contacts are likely not restricted to the new borrowings. There seem to be borrowings from Finnic which, in addition to Komi, are present also in Udmurt. Moreover, there are also words which seem to have been adopted from Pre-Finnic into Proto-Permian (SAARIKIVI 2006: 33–38). Thus it seems that pre-Finnic and pre-Permian language forms have had fairly long lasting and intimate contact. This same observation has even earlier been made by JORMA KOIVULEHTO in connection with early Germanic and other western Uralic borrowings which spread into the Permian languages (KOIVULEHTO 1981; 1989). This view is also supported by the fact that Finnic toponyms in the Dvina basin point to a Finnic substrate language of archaic character which likely spread into the region as soon as Proto-Finnic began to break up.

In addition to analysing the Finnic, Sámi and Permian layers of toponymy, it is a tradition in Russian onomastic studies to distinguish Meryan and *севернофинская* ('North Finnic') layers of substrate toponymy. Both of these layers of toponymy are, according to MATVEEV (1996, 1998, 2001), spread in the southern parts of the Akhangelsk Region.

The central Russian tribe name *мерья* is attested in several historical sources and there are ethnotoponyms from the same word stem. It is hard to define the distinctive Meryan types of toponyms, however, because the *мерья* is just a tribe name in the Chronicles, not a language that would have been described by linguists. Most of the toponymic types present in the territory connected with *мерья* in historical sources are also present elsewhere. Thus,



<sup>64</sup> This view was based, among other things, on the toponym *Тоймокары* which figures in the Chronicles and presumably includes the Permian word *kar* 'fortified place'.

the northern Russian topoformants *-ма*, *-н(ь)га*, *-пола*, *-бала* and *-ла* occur in toponyms also in the territory historically inhabited by Merya. In the same area there are toponymic types with the closest cognates in Mordvinian, such as the river name formant *-ля* (~ Mordvinian *l'ej* 'river') and the formant *марь* (~ Mordvinian *mar* 'hummock', for details see AHLQVIST 2001). The toponymic types explained as Meryan in the south of the Dvina basin (most notable by the rivers Ustja and Vashka) have been even otherwise explained, as a heritage of some groups of Maris (AHLQVIST 1997; 2000). Without going into details, it is sufficient to note that there are parallels between the pre-Slavic toponymy of the southern Dvina basin and the Jaroslavl and Kostroma areas. This is only natural in view of the political dependence of these areas on the central Russian principalities. In order to label a toponymy layer of some region as Meryan, however, one should define which toponymic types should be classified as Meryan. Before this is done, Meryan is not too useful a characterisation for a layer of substrate toponyms.

The *севернофинская* type of toponymy is even less clearly defined. Most of its area falls outside the Arkhangelsk Region and the scope of this presentation. According to MATVEEV, a characteristic feature of this group is the preservation of Uralic \*š (which developed into *s* in Finnic and *č* in Sámi). This would be reflected in those toponyms with the base *селм-* (< \*šolma > Finnic *salmi*, Sámi *čolbmi* 'strait'). Though the characteristics of the *севернофинская* group have never been explicitly presented, the idea that in northern Russia there once existed an archaic Uralic substrate language which did not undergo either Finnic or Sámi sound shifts finds some support, in that some substrate toponyms probably did not undergo the Finnic sound shift *š > h* (cf. toponymic type *ныи-* 'sacred' [> Fi. *pyhä*], for details see MATVEEV 2004: 232–242, cf. also the dialectal word *сорьез* 'grayling' which could correspond to the Finnish *harjus* id. [MYZNIKOV 2003: 75]). Some other toponyms likely preserved word initial \*wo (cf. formant *-вой*, *-бой*) which later developed into *o* in Finnic and Sámi.

### 6.5. Northern Russian toponymy and the origin of Uralic subbranches

Needless to say, ethnic conclusions made on the basis of northern Russian toponyms are uncertain because of the varying reliability of the toponymic etymologies they are based on. Notwithstanding these difficulties, some general remarks can be made.

The Proto-Uralic linguistic homeland was, most likely, situated in the southern taiga zone (ITKONEN 1966; JOKI 1973: 358–364; CARPELAN–PARPOLA 2002).<sup>65</sup> Therefore, one must suppose that also the Arkhangelsk Region was

<sup>65</sup> In the scholarly history, the Uralic linguistic homeland has most often been located either in the southern taiga zone of western Siberia (CASTRÉN, HAJDÚ, JAN-



linguistically non-Uralic at the time it was first settled by humans. It is quite probable that some of the Pre-Uralic toponyms have been preserved in river names. As the Dvina basin is closer to the linguistic homeland of Uralic than the areas in which Finnic and Sámi are spoken at present, it seems likely that at least some parts of this area became linguistically Uralic before the present Finnic and Sámi speaking areas.

As the Permian toponyms in the Dvina basin are of modest number and probably relatively new, the Proto-Permian homeland must have been outside of this territory. This observation is in accordance with the prevailing theories concerning the location of the Proto-Permian speaking area somewhere in the Vjatka basin (BARTENS 2001: 10–11; BELYKH 1999).

The Finnic toponymy of the Dvina basin has at least two and probably more layers. Also, many Germanic and Baltic loanwords (*\*lampas* ‘sheep’ [< Germanic], *\*ranta* ‘shore’ [< Germanic], *\*varkas* ‘thief’ [< Germanic], *\*härkä* ‘ox’ [< Baltic], *\*liiva* ‘sand’ [< Baltic], *\*kelta* ‘yellow’ [< Baltic]) occur in Finnic substrate toponyms of Dvina basin (cf. 6.2. and 6.3. above) and the Finnic substrate languages of the area are thus “modern Finnic”, unlike the Sámi (or whatever they should be labeled) substrate languages which cannot be characterised as “modern Sámi” because of the lack of one central vocabulary layer.

Due to the archaic phonological characteristics of some extinct Finnic dialects of the Dvina basin, the Finnic language must have spread to this area quite early. At present, standard theories locate Proto-Finnic somewhere in the vicinity of the Gulf of Finland (KALLIO 2006 with relevant references). The main reason for this is the Proto-Finnic and even earlier borrowings from Proto-Germanic which must have been adopted somewhere in the vicinity of the Gulf of Finland, as there is no evidence of Germanic tribes in inner Russia.<sup>66</sup> Aside from Germanic loanwords, there are other layers of

---

HUNEN) or in the Middle Volga region (AMINOFF, TOIVONEN, CARPELAN–PARPOLA). At the present, the palaeolinguistic argumentation by CARPELAN–PARPOLA (2002) seems most convincing. As there are established borrowings from Proto-Indo-European in Proto-Uralic, the latter must have been spoken in the vicinity of the former. The Proto-Indo-European homeland, in turn, can be located by cart and wheel vocabulary and the archaeological findings connected with early cart and wheel culture in the Ukrainian steppe (cf. MALLORY 1989). Thus, the Uralic linguistic homeland must have been situated north of this territory, in the Middle Volga region. In addition to loan contacts, this explains the areal distribution of the Uralic languages. It also fits in with the palaeolinguistically meaningful vocabulary reconstructable in Proto-Uralic.



<sup>66</sup> Proto-Germanic loanwords in Finnic have traditionally been connected to archaeologically discernable Bronze Age influences in the western coasts of Finland and

borrowings in Proto-Finnic and Proto-Sámi which point to a more eastern Finnic homeland, however. The Baltic loanwords may have been adopted both in the vicinity of the Gulf of Finland as well as in central European Russia, but it is especially the Iranian borrowings (cf. KOIVULEHTO 1999b) that imply language contacts in central Russia. Further, the borrowings from Proto-Finnic and even earlier western Uralic language forms to Proto-Permian point to an early presence of Finnic tribes surprisingly far away in the east. The Finnic languages seem thus to have formed a dialect continuum in which Germanic loanwords have spread as far as Proto-Permian and, in the later period, Komi. As part of the same dialect continuum Aryan and Iranian loanwords may have spread from central Russia to dialects which later developed into modern Finnic. Also, sound shifts ( $\check{s} > h$ ) have probably spread in this way most likely from west to east (and it has traditionally been argued that Proto-Finnic sound shifts originated through Germanic influence [see POSTI 1953, KALLIO 2000]). This is supported by the fact that those toponyms which probably did not undergo the shift  $\check{s} > h$  are concentrated to the east of the Dvina basin (cf. MATVEEV 2004: 234–242).

In the later period, new Finnic tribes spread from west to east and brought new toponymic models with Karelian phonological characteristics to the north of the Dvina basin. Veps, in turn, spread into the southwest of the Arkhangelsk Region. The old Finnic population of the Dvina basin was neither Karelian nor Veps, however. They seem to have spoken an archaic language with several Proto-Finnic features and, quite probably, one development in common with the southern group of Finnic (mid-central vowel). Thus it seems that the division of the Finnic languages into a southern and a northern group has old roots.<sup>67</sup> The area in which the southern dialects began to emerge was probably situated east of Estonia by Lake Peipus. The spread of an archaic Finnic language form from this area both to the Arkhangelsk Region and to southern Estonia would be understandable.

The present-day Arkhangelsk Region and its neighbouring territories probably played an important role in the development of the Sámi languages as well. As noted above, there are no examples of differentiating Sámi geographical vocabulary in the area whereas the traces of the Sámi sound shifts are likely restricted to the western parts of the area. Moreover, many toponymic types, with the probable Sámi etymologies include lexemes ety-

---

Estonia. One should note, however, that many germanists consider the dispersal of Proto-Germanic as a substantially later phenomenon.

<sup>67</sup> Similar point of view has recently been expressed also by KALLIO (2007), though on different grounds.



mologically opaque in Sámi (saN *čukcá* ‘capercaillie’ *siida* ‘village’, *njukča* ‘swan’, *suhpi* ‘aspen’ and their counterparts<sup>68</sup>).

According to MATVEEV (1999, 2001, 2004), Finnic and Sámi substrate toponyms exist side by side almost everywhere in the Dvina basin. Such a conclusion seems to be an illusion caused by too straightforward an ethnic interpretation of the toponymic material, however. As the Sámi toponymic layer is very different from that of modern Sámi, it is quite possible that many toponymic types characterised as Sámi by MATVEEV originated in idioms closer to Finnic or Pre-Finnic. At the present phase of research it cannot be established whether toponyms such as *Чухчамень* with lexemes etymologisable both on the basis of Sámi and Finnic originated from the same kind of extinct idioms as toponyms with formants *-сапа* and *-нала* or the toponyms with the base *кыч-* characterised as Finnic (although they do not point to any particular living Finnic language) or in substrate languages which were fundamentally different from Finnic. In any case, there are toponymic types which cannot be identified as either Finnic or Sámi.

From the point of view of the identification of substrate languages affricates are of great importance. There are namely certain toponymic types which seem to have preserved the nonpalatised affricate *\*c*, cf. *неч-* ‘spruce’ (< *\*pecä* [> Finnish *petäjä*, North Sámi *beacci*]), *ноч-* ‘branch of a river’ (< ?*\*puca* [> Finnish *pudas*<sup>69</sup>], cf. even the etymologies of MATVEEV *куч-* ‘rotten’ [> saN *guocca*], *кочкем-* ‘eagle’ [> saN *goaskin*]). This affricate seems to have also been preserved in South Estonian (see KALLIO forthcoming), but in the other Finnic languages it has developed into *t* or *s* (latter reflex before *i*). Thus the toponymic types referred to above have, if their etymologies are correct, preserved the Proto-Finnic consonantism and, in this respect, they stand apart from most of the Finnic. Moreover, as noted above, there are examples of a Sámi phonological shift *\*ś* > *\*ć* in the toponyms. In addition, as also noted above, there are also some toponyms which have probably preserved Proto-Uralic *\*š* and word initial *\*wo*.

<sup>68</sup> The two first two of these do not have any cognates in the other Uralic languages. The two latter display phonological irregularities (such as word initial *n* in *njukča* and initial syllable *u* instead of the regular *uo* in *suhpi*) and even the words considered as their cognates have many irregularities (cf. Finnish *haapa* ‘aspen’ with irregular long *a*, Mordvinian *loksij* ‘swan’ with *l* instead of *j*, etc.)

<sup>69</sup> Finnish *pudas* has a regular cognate in Ob-Ugrian languages (mansi *pasəl*, *posəl*, *posal* Khanty *pāsəl*, etc. ‘river branch’). The Proto-Uralic form of the word would be *\*puca*. It is quite probable that the North Russian toponyms with the base *ноч-* belong here as many of them denote river branches. In this case, the phonetic form of the word is quite interesting, with a preserved back affricate and a vocalism close to Sámi (note, however, that saN *bovces* ‘river branch’ does not belong here because of the *-vc-*).



Thus, there seem to be remnants of archaic language forms with a consonantism close to Pre-Finnic (or Proto-Uralic as these are almost identical at the reconstruction level) in the Arkhangelsk Region and neighbouring areas. From the point of view of linguistic prehistory this would be only natural: as the inland area west and northwest of the Uralic linguistic homeland must have become linguistically Uralic before the Baltic Sea coast, where the (Pre-)Finnic-Germanic language contacts presumable took place, it is necessary to assume that those languages which first spread to this area were of a phonologically archaic character. While the Finnic language form spread to these areas from the west some enclaves of these archaic Uralic language forms seem to have escaped this second wave of Uralicisation and probably survived until the Slavicisation of the area.

The Proto-Sámi sound shifts seem to have originated in that area which later became Finnic. After  $*ś > *ć$ , a change which probably occurred in the common ancestor of Proto-Finnic and Proto-Sámi, Sámi vowel rotation ( $*a > uo, i, *e > a, ea, *ä > á, ie$ , etc.) took place. The Sámi vowel changes are, quite probably, attested in toponyms in the western parts of the Arkhangelsk Region. As Proto-Sámi also had multiple contacts with Proto-Germanic (cf. KOIVULEHTO 2000; AIKIO 2006), it can be assumed that in a similar manner to the Finnic dialect continuum described above, there was also a Sámi dialect continuum capable of spreading Germanic loanwords from the, what is nowadays, the Finnish coast of the Baltic Sea to the east. The area of the Sámi languages must have been situated to the north and probably also to the east of the Finnic dialect continuum. In the area west of the Arkhangelsk Region Proto-Sámi speakers also encountered populations who spoke a Palaeo-European language(s), from whom they borrowed vocabulary that did not spread into the Dvina basin.

As the Sámi lexemes present in the toponyms of the Arkhangelsk Region are largely opaque in that they do not represent regular Sámi sound shifts, one is inclined to conclude that the rare lexical parallels between the toponymy of the Pinega basin and the Sámi languages may be due to borrowing. For example, the toponymic base *ууѳ-* which MATVEEV associates with the Proto-Sámi  $*supē$  ‘aspen’, appears over a large area in which the prevailing toponymic substrate type is Finnic (MATVEEV 2004: 318). Moreover, this word also appears in toponyms which have distinctively Finnic bases and formants (*Шуѳматка, Шуѳоя*). In the same area, toponyms formed from the Finnic  $*haapa$  ‘aspen’ do not exist (MATVEEV 2004: 308, 318). Thus one could imagine that the Finnic idioms of the Pinega and neighbouring dialects might have borrowed the designation of aspen from the Proto-Sámi found at that time in the western parts of the present-day Arkhangelsk Region. This word would then have become commonplace in the Finnic



toponyms of this area. As for toponymic types such as *чухч*- ‘capercaillie’ and *нюхч*- ‘swan’ (cf. also *шуд*- ‘winter village, SAARIKIVI 2004b: 211) it seems premature to make a suggestion concerning what the mechanism was for their diffusion to the east. Probably, some of these words may be Palaeo-European substrate loans borrowed by Proto-Sámi speakers either from a Uralic speaking population in the Dvina basin or from their non-Uralic speaking predecessors.

### References

- AGEEVA, R. A. (1989) • Агеева, Р. А. *Гидронимия Русского Северо-Запада как источник культурно-исторической информации*. АН СССР. Институт научной информации по общественным наукам. Москва, Наука.
- AHLQVIST, AUGUST (1887) *Kalevalan karjalaisuus. Kalevalasta itsestään ja muualta*. Todistanut AUG. AHLQVIST. Helsinki, Suomen yliopisto.
- AHLQVIST, ARJA (1997) • Альквист, Арья *Мерьянская проблема на фоне многослойности субстратной топонимии*. In: *Вопросы языкознания* 1997/6, p. 2–36.
- AHLQVIST, ARJA (2000) • Альквист, Арья *Мерьяне, не мерьяне*. In: *Вопросы языкознания* 2000/2, p. 45–34, 2000/3, p. 83–97.
- AHLQVIST, ARJA (2001) • Альквист, Арья *Субстратная топонимия Ярославского Поволжья*. In: *Очерки исторической географии. Северо-Запад России. Славяне и финны*. Ed. by ГЕРД, А. С.—ЛЕБЕДЕЕВА, Г. С. Санкт-Петербург, Издательство Санкт-Петербургского университета, p. 436–467.
- AIKIO, ANTE (2001) *Sápmelas loatnasánit suomelas gielaid davvejoavkkus. Pro gradu-dutkamus sámegielas*. Oulu universitehta. Manuscript.
- AIKIO, ANTE (2004) *An essay on substrate studies and the origin of Sámi*. In: *Etymologie, Entlehnungen und Entwicklungen. Festschrift für Jorma Koivulehto zum 70. Geburtstag*. Ed. by HYVÄRINEN, IRMA—KALLIO, PETRI—KORHONEN, JARMO. Mémoires de la Société Neophilologique de Helsinki 63, p. 5–34.
- AIKIO, ANTE 2006 *On Germanic-Sámi contacts and Sámi prehistory*. JSFOu. 91. p. 9–55.
- AIKIO, ANTE & ASLAK (2001) *Heimovaelluksista jatkuvuuteen — suomalaisen väestöhistorian tutkimuksen pirstoutuminen*. Muinaistutkija 2001/4, p. 2–21.
- AIKIO, SAMULI (2003) *Inarin ja inarilaisten nimiä*. In: *Inari — Aanaar. Inarin historia jääkaudesta nykypäivään*. Ed. by LEHTOLA, VELI-PEKKA. Inarin kunta, p. 96–113.



- AINIALA, TERHI (1997) *Muuttuva paikannimistö*. SKST 667.
- АТАМАНОВ, М. Г. (1988) • Атаманов, М. Г. *Удмуртская ономастика*. Ижевск. Удмуртия.
- BARTENS, RAJA (2001) *Permiläisten kielten rakenne ja kehitys*. MSFOu. 238.
- BELYKH, S. K. (1999) • Белых, С. К. *К вопросы локализации прародины пермян. Пермский мир в раннем средневековье*. РАН, Уральское отделение, Удмуртский ИЯЛИ. Ижевск.
- BERNSHTAM, T. A. (1973) • Бернштам, Т. А. *Роль верхневолжской колонизации в освоении Русского Севера. (IX–XV вв.)*. Фолклор и этнография Русского Севера. Ленинград.
- CARPELAN, CHRISTIAN (1999) *Käännekohtia Suomen esihistoriassa aikavälillä 5100–1000 eKr.* In: *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*. Ed. by FOGELBERG, P. Bidrag till kännedom av Finlands natur och folk 153. Helsinki, Societas Scientiarum Fennicae, p. 249–280.
- CARPELAN, CHRISTIAN—PARPOLA, ASKO (2002) *Emergence, Contacts and Dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in Archaeological Perspective*. In: *Early Contacts between Uralic and Indo-European: Linguistic and archaeological considerations*. Ed. by CARPELAN, CHRISTIAN—PARPOLA, ASKO—KOSKIKALLIO, PETTERI. MSFOu. 242. Helsinki, p. 55–150.
- CASTRÉN, M. A. (1844) *Anmärkningar om Sawolotschesskaja Tshud*. Suomi 1844, p. 3–22.
- DMITRIEVA, T. N. (2006) • Дмитриева, Т. Н. *Топонимия реки Казым*. Екатеринбург. УрГУ.
- EUROPAEUS, D. E. D. (1868–1870) *Tietoja suomalais-ungarilaisten kansojen muinaisista olopaikoista*. Suomi II. 7, p. 1–190. 8, p. 27–106.
- FASTER, MARIKO—SAAR, EVAR *Võromaa kotussõnumist*. Ed. By PÄLL, PEETER. Võru instituudi toimõndusõq 13. Võro, Võro instituut.
- GALKIN, I. S.—VORONTSOVA, O. P. (2002) • Галкин, И. С.—Воронцова, О. П. *Топонимия республики Марий Эл. Историко-этимологический анализ*. Йошкар-Ола, Издательство марийского полиграфкомбината.
- GLINSKIKH, G. V. (1983) • Глинских, Г. В. *О принципах и методах исследования субстратной топонимии*. In: *Вопросы ономастики 17. Методы топонимических исследований. Сборник научных трудов*. Свердловск, Издательство УрГУ., p. 3–22.
- GRÜNTAL, RIHO (1997) *Livvistä liiviin. Itämerensuomalaiset etnonymit*. Castreanimumin toimitteita 53. Helsinki.
- GUSENIKOVA, M. L. (1994) • Гусельникова, М. Л. *Полукальки в топонимии Русского Севера*. Канд. дисс. Екатеринбург, УрГУ. Manuscript.
- HAAVIO, MARTTI (1965) *Vjarmien vallan kukoistus ja tuho. Historiaa ja runoutta*. Runot suomentanut Aale Tynni. Porvoo–Helsinki, WSOY.



- NAKULINEN, LAURI (1979) *Suomen kielen rakenne ja kehitys*. Otavan korkeakoulukirjasto. Helsinki, Otava.
- HAUSENBERG, ANU-REET (1983) *Probleme der ostseefinnisch-permischen Sprachkontakte*. Akademie der Wissenschaften der Estnischen SSR. Preprint KKI 23.
- HELIMSKIJ, J. A. (1986) • Хелимский, Е. А. *О прибалтийско-финском языковом материале в новгородских берестяных грамотах*. In: *Новгородские грамоты на бересте (из раскопок 1977–1983 гг.)*. VIII. Ed. by ЯНИН, В. Л.—ЗАЛИЗНЯК, А. А. Москва, p. 252–259.
- HOLST, JAN HENRIK (2001) *Die Herkunft des Estnischen Vokals õ*. In: *JSFOu*. 89, p. 57–98.
- HÄKKINEN, KAISA (1985) *Suomen kielen sanaston historiallista taustaa*. Fennistica 7. Turku, Åbo Akademi.
- ITKONEN, ERKKI (1966) *Suomalais-ugrilaisen kielen- ja historian tutkimuksen alalta*. Tietolipas 20. Helsinki, SKS.
- ITKONEN, T. I. (1948) *Suomen lappalaiset vuoteen 1945. Ensimmäinen osa*. Porvoo–Helsinki. WSOY.
- ITKONEN, TERHO (1998) *Reflections on Pre-Uralic and Sámi-Finnic protolanguage*. In: *FUF* 54, p. 229–266.
- JACKSON, T. N. (1993) • Джаксон, Т. Н. *Исландские королевские саги о Восточной Европе (с древнейших времен до 1000 г.)*. Тексты, перевод, комментарии. Древнейшие источники по истории народов восточной Европы. Москва, Наука.
- JANHUNEN, JUHA (1977) *Samojedischer Wortschatz. Gemeinsamojedische Etymologien*. Castrenianumin toimitteita 17. Helsinki.
- JOKI, AULIS J. (1973) *Uralier und Indogermanen. Die älteren Berührungen zwischen Uralischen und Indogermanischen Sprachen*. MSFOu. 151. Helsinki.
- КАВИНИНА, N. V. (1997) • Кабинина, Н. В. *Топонимия дельты Северной Двины*. Канд. дисс. Екатеринбург, УрГУ. Manuscript.
- KALIMA, JALO (1919) *Die ostseefinnischen Lehnwörter der Russischen*. MSFOu. 44.
- KALIMA, JALO (1944a) *Zu den anlautenden Sibilanten der ostseefinnischen Ortsnamen im Russischen*. In: *FUF* 28, p. 151–162.
- KALIMA, JALO (1944b) *Einige Russische Ortsnamentypen*. In: *FUF* 28, p. 99–150.
- KALIMA, JALO (1946) *Venäjän vierasperäisten nimien tutkimisesta*. Suomalaisen Tiedeakatemia esitelmät ja pöytäkirjat 1945. Helsinki, Suomalainen tiedeakatemia, p. 101–110.
- KALLASMAA, MARJA (2000) *Saaremaa kohanimed 2*. Tallinn, Eesti keele instituut.



- KALLASMAA, MARJA (2003) *Läänemurde loodus- ja viljelusnimed*. Ed. by ROSS, EEVI. Tallinn, Eesti keele sihtasutus.
- KALLIO, PETRI (2000) *Posti's superstrate theory at the threshold of the new millenium*. In: *Facing Finnic. Some challenges to historical and contact linguistics*. Ed. by LAAKSO, JOHANNA. Castreanum in toimitteita 59, p. 80–99.
- KALLIO, PETRI (2006) *Suomen kantakielten absoluuttista kronologiaa*. In: *Virittäjä* 110, p. 2–25.
- KALLIO, PETRI (2007) *Kantasuomen konsonanttihistoriaa*. In: *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*. Doaim. JUSSI YLIKOSKI ja ANTE AIKIO. MSFOu 253. p. 229–250.
- KANNISTO, ARTTURI (1927) *Über die früheren Wohngebiete der Wogulen im Lichte der Ortsnamenforschung*. In: *FUF* 18, p. 57–89.
- KESKJA = ЛЫТКИН, В. И.—ГУЛЯЕВ, Е. С. *Краткий этимологический словарь коми языка. Переиздание с дополнением*. Сыктывкар, Коми книжное издательство. 1999.
- KETTUNEN, LAURI (1954–1955) *Etymologische Untersuchung über estnische Ortsnamen*. *Annales Academiae Scientiarum Fennicae*. B 90, p. 1–2.
- KIRKINEN, HEIKKI (1963) *Karjala idän kulttuuripiirissä. Bysantin ja Venäjän yhteyksiä keskiajan Karjalaan*. Helsinki, Kirjayhtymä.
- KIVINIEMI, EERO (1971) *Suomen partisiippinimistöä. Ensimmäisen partisiippin sisältävät henkilö- ja paikannimet*. SKST 295.
- KIVINIEMI, EERO (1977) *Väärät vedet. Tutkimus mallien osuudesta nimistönmuodostuksessa*. SKST 377.
- KIVINIEMI, EERO (1990) *Perustietoa paikannimistä*. Suomi 149. Helsinki, SKS.
- KOIVULEHTO, JORMA (1981) *Paikan ja joukon tulo kieleen*. In: *Virittäjä* 85, p. 195–213.
- KOIVULEHTO, JORMA (1989) *Ehkä ja ehto, yskä ja ystävä*. In: *JSFOu*. 82, p. 171–192.
- KOIVULEHTO, JORMA (1997) *Were the Baltic Finns "clubmen"?* In: *You name it. Perspectives on onomastic research*. Ed. by PITKÄNEN, RITVA LIISA—MALLAT, KAIJA. *Studia Fennica linguistica* 7, p. 151–169.
- KOIVULEHTO, JORMA (1999a) *Das Verhältnis des Ostseefinnischen und Lappischen im Lichte der älteren Lehnwörter: die Substitution des fremden Wortausgangs \*-CVz im Lappischen*. In: *Sprachen in Finnland und Estland*. Ed. by LEHTIMÄKI, PEKKA. Wiesbaden, p. 7–22.
- KOIVULEHTO, JORMA (1999b) *Varhaiset indoeurooppalaiskontaktit. Aika ja paikka lainasanojen valossa*. In: *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*. Ed. by FOGELBERG, PAUL. *Bidrag till kännedom av Finlands nature och folk*. 131. Helsinki, Societas Scientiarum Fennicae, p. 207–236.



- KOIVULEHTO, JORMA (2000) *Contact with non-Germanic languages II. Relations to the east*. In: *The Nordic languages. An international handbook of the history of the North Germanic languages*. Vol 1. Ed. by BUNDLE, OSKAR—BRAUNMÜLLER, KURT—JAHR, ERNST HÅKAN—KRAKER, ALLAN—NAUMANN, HANS-PETER—TELEMAN, ULF. Walter de Gruyter, Berlin—New York, p. 583–594.
- KOLPAKOV, E. M.—RYABTSEVA, E. N. (1994) *A new type of Chud burial construction*. *Fennoscandia Archaeologica* 11. Helsinki, Suomen arkeologinen seura, p. 77–86.
- КОМУЯГИНА, Л. Р. (1994) • Комягина, Л. П. *Лексический атлас Архангельской области*. Архангельск, Издательство международного университета им. М. В. Ломоносова.
- KORHONEN, MIKKO (1981) *Johdatus lapin kielen historiaan*. SKST 370.
- KOSKI, MAUNO (1967–1970) *Itämerensuomalaisten kielten hiisi-sanue: semanttinen tutkimus*. Turun yliopiston julkaisuja. Sarja C. Scripta lingua fennica edita. 5, 7.
- KUZ'MIN, D. V. (2004) • Кузьмин, Д. В. *Ареальная дистрибуция топонимных моделей Беломорской Карелии*. Канд. дисс. Петрозаводск. Manuscript.
- LASHUK, L. P. (1958) • Лашук, Л. П. *Очерк этнической истории Печорского края. Опыт историко-этнографического исследования*. Сыктывкар, Коми книжное издательство.
- LEHTIRANTA, JUHANI (1989) *Yhteissaamelainen sanasto*. MSFOu. 200.
- LYTKIN, VASILI (1967) *Vepsäläiset ja komisyryjäanit*. In: *Virittäjä* 71, p. 23–27.
- МАКАРОВ, Н. А. (1993) • Макаров, Н. А. *Русский Север. Таинственное средневековье*. Москва, Институт археологии РАН — Российский фонд фундаментальных исследований.
- МАКАРОВ, Н. А. (1997) • Макаров, Н. А. *Колонизация северных окраин Древней Руси в XI–XIII вв. По материалам археологических памятников на волоках Белоозерья и Поонежья*. Москва, Скрипторий.
- MALLORY, JOSEPH (1989) *In search of the Indo-Europeans. Language, archaeology and myth*. London, Thames & Hudson.
- МАНЮКИН, И. С. (2002) • Манюхин, И. С. *Происхождение саамов. Опыт комплексного изучения*. Петрозаводск, Карельский научный центр РАН, ИЯЛИ.
- МАТВЕЕВ, А. К. (1978) • Матвеев, А. К. *Русская диалектная чильма. Этимология русских диалектных слов*. Свердловск, УрГУ, p. 54–59.
- МАТВЕЕВ, А. К. (1980) • Матвеев, А. К. *Этимологический словарь субстратных топонимов Русского Севера. Том 1. Прибалтийско-финские и саамские форманты и основы*. Свердловск, УрГУ. Manuscript.



- МАТВЕЕВ, А. К. (1986) • Матвеев, А. К. *Методы топонимических исследований*. Учебное пособие. Свердловск, УрГУ.
- МАТВЕЕВ, А. К. (1996) • Матвеев, А. К. *Субстратная топонимия Русского Севера и мерянская проблема*. In: *Вопросы языкознания* 1996/1, p. 3–24.
- МАТВЕЕВ, А. К. (1998) • Матвеев, А. К. *Мерянская топонимия на Русском Севере. Фантом или феномен?* In: *Вопросы языкознания* 1998/5, p. 90–105.
- МАТВЕЕВ, А. К. (1999) • Матвеев, А. К. *Древнее население севера европейской части России. Опыт лингвоэтнической карты 1*. Известия Уральского государственного университета. Гуманитарные науки, Вып. 2. 13/1999. Екатеринбург, УрГУ, p. 80–88.
- МАТВЕЕВ, А. К. (2001) • Матвеев, А. К. *Субстратная топонимия Русского Севера 1*. Екатеринбург, Издательство Уральского университета.
- МАТВЕЕВ, А. К. (2002) • Матвеев, А. К. *Субстратные лимнонимы Русского Севера*. Финно-угорское наследие в русском языке 2. Екатеринбург, УрГУ, p. 190–194.
- МАТВЕЕВ, А. К. (2004) • Матвеев, А. К. *Субстратная топонимия Русского Севера 2*. Екатеринбург, Издательство Уральского университета.
- МЕЛНИКОВА, Y. A. (1986) • Мельникова, Е. А. *Древнескандинавские географические сочинения. Тексты, перевод, комментарии*. In: *Древнейшие источники по истории народов восточной Европы*. Ed. by ЯНИН, В. Л. Москва, Наука.
- MIKKOLA, J. J. (1894) *Die ältere Berührungen zwischen Ostseefinnisch und Slavisch. 1. Slavische Lehnwörter in den westfinnischen Sprachen*. MSFOu. 8.
- МИКНР = *Материалы по истории крестьянского хозяйства и подвинностей XVI–XVII вв.* Ed. by ВОСКБОЙНИКОВА, Н. П.—КОПАНЕВ, А. И.—ПОБЕДИМОВА, Г. А. Москва–Ленинград, Академия Наук СССР, Ленинградское отделение института истории СССР. Ленинградское отделение археологической комиссии. 1977.
- MULLONEN, I. I. (1988) • Муллонен, И. И. *Гидронимия бассейна реки Ояти*. Петрозаводск, Карелия.
- MULLONEN, I. I. (1994) • Муллонен, И. И. *Очерки вепсской топонимии*. РАН, Карельский научный центр, ИЯЛИ. Санкт-Петербург, Наука.
- MULLONEN, I. I. (2002) • Муллонен, И. И. *Топонимия Присвирья. Проблемы этноязыкового контактирования*. Петрозаводск, РАН, Карельский научный центр, ИЯЛИ.
- MUZNIKOV, S. A. (2003) • Мызников, С. А. *Атлас субстратной и заимствованной лексики русских говоров Северо-Запада*. Санкт-Петербург, Наука.



- MYZNIKOV, S. A. (2004) • Мызников, С. А. *Лексика финно-угорского происхождения русских говоров Северо-Запада. Этимологический и лингвогеографический анализ.* Санкт-Петербург, Наука.
- NASONOV, A. N. (1951) • Насонов, А. Н. “Русская земля” и образование территории Древнерусского государства. Москва, Наука.
- NAZARENKO, V. A.—OVSYANNIKOV, O. V.—RYABININ, Y. A. (1984) • Назаренко, В. А.—Овсянников, О. В.—Рябинин, Е. А. *Средневековые памятники Чуди заволочской.* In: *Советская археология* 1984/4, p. 197–215.
- NUUTINEN, OLLI (1989) Järvi — *baltilainen laina.* In: *Virittäjä* 93, p. 497–503.
- OVSYANNIKOV, O. V. (1980) *First-discovered Burialfield of “Zavolochye Tschud”.* Fenno-ugri et slavi. Symposium: “The Cultural Relations between the Peoples and the Countries of the Baltic Area during the Iron Age and the Early Middle Ages”. Helsingin yliopiston arkeologian laitos. Moniste 22. Helsinki, p. 228–236.
- РОПОВ, А. И. (1965) • Попов, А. И. *Географические названия. Введение в топонимику.* АН СССР, Географическое общество СССР. Москва–Ленинград, Наука.
- РОПОВА, Y. Y. (1999) • Попова, Е. Ю. *Этнонимия Русского Севера.* Канд. дисс. УрГУ, Екатеринбург. Manuscript.
- POSTI, LAURI (1953–1954) *From Pre-Finnic to late Proto-Finnic.* In: *FUF* 31, p. 1–91.
- POZHARITSKAYA, S. K. (1997) • Пожарицкая, С. К. *Русская диалектология.* Москва. МГУ, Издательство Московского университета.
- RYABININ, Y. A. (1997) • Рябинин, Е. А. *Финно-угорские племена в составе Древней Руси. К истории славяно-финских этнокультурных связей. Историко-археологические очерки.* Санкт-Петербургский государственный университет — Институт истории материальной культуры РАН. Санкт-Петербург, Издательство Санкт-петербургского университета.
- REW = MAX VASMER *Russisches etymologisches Wörterbuch.* Indogermanische Bibliothek, 2. Reihe (Wörterbücher). Heidelberg, Carl Winter, 1953–1958.
- RÄISÄNEN, ALPO (2003) *Nimet mieltä kiehtovat. Etymologista nimistöntutkimusta.* SKST 936.
- SAARIKIVI, JANNE (2000) *Kontaktikähtöinen kielenmuutos, substraatti ja substraattinimistö.* In: *Virittäjä* 104, p. 392–415.
- SAARIKIVI, JANNE (2002) • Саарикиви, Я. *Саамская субстратная топонимия Русского Севера с точки зрения сравнительной финно-угристики.* In: *Финно-угорское наследие в русском языке* 2. Екатеринбург, УрГУ, p. 94–112.



- SAARIKIVI, JANNE (2003) • Саарикиви, Я. Прибалтийско-финская антропонимия в субстратных топонимах Русского Севера. Перспективы изучения. In: *Этимологические исследования* 8. Екатеринбург, УрГУ, p. 136–147.
- SAARIKIVI, JANNE (2004a) *Is there Palaeo-European substratum interferences in the western branches of Uralic?* In: *JSFOu*. 90, p. 187–214.
- SAARIKIVI, JANNE (2004b) *Über die saamischen Substratennamen des Nordrusslands und Finnlands*. In: *FUF* 58, p. 162–234.
- SAARIKIVI, JANNE (2004c) *Namenforschung am Swir*. In: *FUF* 58, p. 344–352.
- SAARIKIVI, JANNE (2006) *Substrata Uralica. Studies on Finno-Ugrian Substrate in Northern Russian dialects. Academic dissertation*. Tartu.
- SAARIKIVI, JANNE—GRÜNTAL, RIHO (2005) *Itämerensuomen uralilainen tausta*. In: *Muuttuva muoto. Kirjoituksia Tapani Lehtisen 60-vuotispäivän kunniaksi*. Ed. by GRÜNTAL, RIHO—SUUTARI, TONI—VAATTOVAARA, JOHANNA. Kieli 16. Helsinki, Helsingin yliopiston suomen kielen laitos, p. 111–146.
- SAMMALLAHTI, PEKKA (1988) *Historical phonology of Uralic languages. With special reference to Samoyed, Ugric and Permic*. In: *Uralic languages. Description, history and foreign influences*. Ed. by SINOR, DENIS. Leiden–New York–København–Köln., E. J. Brill, p. 478–554.
- SARHIMAA, ANNELI (1995) *Karjalan kansat ja kielet kontaktissa*. In: *Virittäjä* 99, p. 191–223.
- SAVELJEVA, E. A. (1992) *Die ethnokulturellen Kontakte der Perm Wytschegodskaja mit den westlichen Nachbarnsvölker*. In: *Suomen varhaishistoria*. JULKU, KYÖSTI. *Studia Historica Septentrionalia* 21. Rovaniemi, Pohjois-Suomen historiallinen yhdistys, p. 495–506.
- SHILOV, A. L. (1999) • Шилов, А. Л. *Заметки по исторической топонимике Русского Севера*. Москва.
- SIMINA, G. Y. (1980) • Сими́на, Г. Я. *Географические названия. По материалам письменных памятников и современной топонимики Пинежья*. АН СССР. Географическое общество СССР. Ленинград, Наука.
- SJÖGREN, ANDREAS JOHANN (1832a) *Ueber die älteren Wohnsitze der Jemen. Ein Beitrag zur Geschichte der Tshudischen Völker Russlands*. Mémoires de l'académie impériale des sciences de St. Pétersbourg, VIeme Série. Sciences politiques, histoire et philologie. St. Pétersbourg, p. 263–346.
- SJÖGREN, ANDREAS JOHANN (1832b) *Wann und wie wurden Sawolotschje und Sawolokschen Tschuden russisch? Ein kritischer Versuch zur Aufklärung der Geschichte des russischen Nordens*. Mémoires de l'académie impériale des sciences de St. Pétersbourg, VIeme Série. Sciences politiques, histoire et philologie. St. Pétersbourg, p. 491–526.



- SNK = MIKKONEN, PIRJO—PAIKKALA, SIRKKA *Sukunimet*. Revised edition. Otava, 2000.
- SRNG = *Словарь русских народных говоров*. Ed. by ФИЛИН, Ф. П. (т. 1–23), СОРОКОЛЕТОВ, Ф. П. (24–). Москва–Ленинград (Санкт-Петербург) 1965–.
- SSA = *Suomen sanojen alkuperä. Etymologinen sanakirja*. Ed. by ITKONEN, ERKKI—KULONEN, ULLA-MAIIJA. SKST 556, Kotimaisten kielten tutkimuskeskuksen julkaisuja 62. 1992–2000.
- СТОЕВКЕ, DETLEF-ЕСКНАРД (1964) *Die ostseefinnischen Personennamen in Rahmen des urfinnischen Namensystems*. Nord- und ost-europäische Geschichtstudien Hamburg. Leibniz.
- ТКАЧЕНКО, О. В. (1985) • Ткаченко, О. В. *Мерянский язык*. АН Украинской ССР. Киев, Наукова думка.
- TURKIN, A. I. (1971) • Туркин, А. И. *Топонимия нижней Вычегды*. Канд. дисс. АН СССР, Институт языкознания. Москва. Manuscript.
- TURKIN, A. I. (1989) = Туркин, А. И. *Топонимия Коми АССР*. Докт. дисс. Москва. Manuscript.
- UINO, PIRJO (1997) *Ancient Karelia: Archaeological studies*. Suomen muinaismuistoyhdistyksen aikakauskirja 104. Helsinki.
- VAHTOLA, JOUKO (1980) *Tornionjoki- ja Kemijokilaakson asutuksen synty. Nimistötieteellinen ja historiallinen tutkimus*. Studia Historica Septentrionalia 3. Rovaniemi, Pohjois-Suomen historiallinen yhdistys.
- VASMER, MAX (1934–36) *Beiträge zur historischen Völkerkunde Osteuropas*. II: *Die ehemalige Ausbreitung der Westfinnen in den heutigen slawischen Länder* (1934). III: *Merja und Tsceremissen* (1935). IV: *Die ehemalige Ausbreitung der Lappen und Permier in Nordrussland*. Berlin: Preussische Akademie der Wissenschaften. (Sonderausgabe aus den Sitzungsberichten der Preussischen Akademie der Wissenschaften. Phil-Hist. Klasse).
- VASMER, MAX (1941) *Die alten Bevölkerungsverhältnisse Russlands im Lichte der Sprachforschung*. Preussische Akademie der Wissenschaften. Vorträge und Schriften. Heft 5. Berlin.
- VEENKER, WOLFGANG (1967) *Die Frage des finnougriischen Substrats in der russischen Sprache*. Indiana University publications, Uralic and Altaic Series 82. Bloomington, Indiana University.
- WIKLUND, K.B (1911) *Lapparnas forna utbredning i Finland och Ryssland belyst af ortnamnen*. In: *Le Monde Oriental* 5, p. 101–136, 175–196.
- VOSTOKOV, A. KH. (1812) • Востоков, А. Х. *Задача любителям этимологии*. Санктпетербургский вестник 2, p. 204–215.
- VOSTRIKOV, O. V. (1990) • Востриков, О. В. *Финно-угорский субстрат в русском языке*. Учебное пособие по спецкурсу. Свердловск, УрГУ.



- ZALIZNYAK, A. A. (2004) • Зализняк, А. А. *Древненовгородский диалект. Второе издание, переработанное с учетом материала находок 1995–2003 гг.* Москва, Языки славянской культуры.
- ZHEREBTSOV, L. N. (1982) • Жеребцов, Л. Н.: *Историко-культурные взаимоотношения коми с соседними народами: X — начало XX вв.* Коми филиал АН СССР, ИЯЛИ. Москва, Наука.

#### **Archive materials**

- NA = Nimiarkisto. Yleiskokoelma. Kotimaisten kielten tutkimuskeskus, Helsinki.
- STE = Севернорусская топонимическая экспедиция. Основная и районная картотека. Уральский государственный университет. Кафедра русского языка и общего языкознания. Лаборатория топонимии. Екатеринбург.