11 TOWARDS A DIGITAL INFRASTRUCTURE FOR KILDIN SAAMI [1]

Michael Rießler

> Finnish, as a language of education and literature,
> would not produce other than ABC-literature.
> If you believe in the possibility of a Finnish literature,
> then you would even believe in the possibility of the
> foundation of an Estonian and a Saami nation and literature.
> (*Nervander 1845, cit. Kuusi-etal. 1983: 9*)

Introduction

The introductory quote is taken from a letter written originally in Swedish by the Swedish-speaking scientist from Finland Johan Jakob Nervander to Johan Vilhelm Snellman, a *fennoman* philosopher and statesman who was central in the establishment of Finnish as the second national language in Finland beside Swedish. The quote nicely illustrates several points relevant to the discussion in the present paper. As we all know, Estonia eventually became a nation state and Finnish and Estonian are in fact well-established languages of education and literature. As national languages they function similarly to Swedish today. Even the Saami are catching up on their Northern European neighbours' nation building processes. In a similar way to other nations, Saami construct their ethnicity and nationhood on the basis of land, tradition, history, culture, etc. Ethnic and national symbolism is well-illustrated in pan-Saami discourses on, for instance, *Sápmi* (North Saami designation for the Saami 'homeland'), *duodji* (North Saami designation for Saami 'traditional' (artisan) handicraft), *joiking* and *reindeer herding*, or national institutions leading these discourses (e.g. ethnic Saami museums in all Nordic countries or the ethnic Saami University College in Guovdageaidnu).

Thus, Nervander's disbelief in the 'national capabilities' of Finns, Estonians and Saami was proven wrong by history. However, the Saami case is different because there is no Saami nation state. Members of the Saami nation are united under one flag and one national anthem, they celebrate February 6 as national day for all Saami and feel represented by the pan-Saami Parliamentary Conference uniting Saami parliamentarians from four countries. Nonetheless, the perhaps 150,000 Saami are scattered as mini minorities over territories belonging to the four northern European nation states, Finland, Norway, Sweden and Russia. There is no single national language for

all Saami, either. Instead, the ten living Saami languages form a group of very closely related but not mutually comprehensible idioms. Six of them have standardised written forms based on different orthographic principles.

Nevertheless, today's Saami languages are continuously transforming from media of mere 'ABC-literature' to true media of education, of literature and of everyday digital communication. The most developed Saami language is North Saami, spoken in a vast area across the northernmost parts of Finland, Norway and Sweden by about 16,000 (Aikio 2003: 35) speakers. Besides being the primary language of everyday communication in several municipalities with a strong proportion of North Saami inhabitants (especially in Kárášjohka/Karasjok and Guovdageaidnu/Kautokeino in Norway as well as in Uhcejohka/Utsjoki in Finland), today North Saami is also used as a medium for education in preschools, primary and secondary schools, and even in professional and academic education. It is used in fine arts and popular culture, in official documents at the local, national and even European administrative levels, it is the language of political debates and, last but not least, of different kinds of printed, broadcast and internet based mass media.

This paper describes one of the more minor Saami languages, Kildin Saami, which is only starting to become a language of education, literature and written communication.

Current language situation of Kildin Saami

Kildin Saami [2] (Endonym *kīllt sām' kīll'*) belongs, along with Ter Saami, to the peninsula group of the East Saamic languages (Sammallahti 1998: 26–34). Originally, Kildin Saami was spoken all over the central inland parts and the central coastal parts of the Kola peninsula. Four dialects of Kildin Saami are still maintained: the Kīllt dialect, the Koarrdegk dialect, the Lujavv'r dialect and the Arsjogk dialect. The neighbouring Saami dialects in the Northwest belong to the Skolt Saami language. Ter Saami dialects were formerly spoken in the eastern parts of the peninsula, but there are practically no Ter speakers left in these areas today. The third neighbouring Saami language, Akkala Saami, was originally spoken in the Southwest of the Kildin Saami dialect area, and has repeatedly been reported to be extinct since 2003 (Rantala et al. 2009), but there are likely Akkala Saami who speak the language at least passively (Elisabeth Scheller p. c.). Since the Kola peninsula is the traditional homeland of the four Saami languages of Russia, they are normally referred to as Kola Saami.

Kildin Saami is the most vital and most developed Saami language in Russia today. The densest Kildin Saami settlement is found in the village and municipal centre Lovozero (Kildin *Lujavv'r*) with 700–800 ethnic Kildin Saami among a total village population of approximately 3,000. Other Kildin speakers live spread over all parts of the Kola peninsula, both in rural and urban settlements, one of them being the
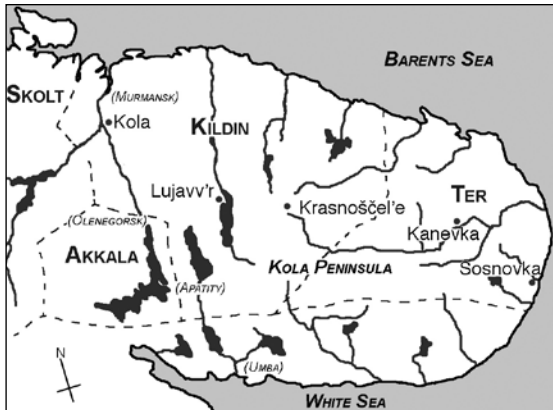
|24| Map of Kola peninsula

administrative centre of the Murmansk area. There is also a considerable Kildin Saami diaspora population, with several speakers among them, who live elsewhere in Russia or in other countries.

Kildin Saami is critically endangered as the result of a very intense language shift to Russian. Today, the language has only about 100 active and perhaps 600 passive speakers (Scheller 2011: 88, 89, 91).

Active language use here refers to fully proficient speakers who use the language on a daily basis and in communications which are not restricted to special thematic domains. There are also fully proficient speakers among the group of passive language users who do not, however, use the language regularly today. The language proficiency of other passive speakers, in Scheller's terminology, might even be restricted to a few expressions or symbolic language use.

|25| Training course on *Archives, Technology and Tools for Kola Saami* (19.–22. October 2011) at the computer lab of the vocational school PU-26 in Lovozero. In the foreground participants Svetlana Danilova (left) and Mariia Medvedeva (right) with the teacher Jeremy Bradley, in the background participants Viktor Danilov and Elisabeth Scheller.

All Kildin Saami speakers are bilingual in Russian. In fact, Russian is the dominant language in all domains of Kildin Saami society, and Kildin Saami is hardly ever heard in public life. Currently, there are hardly any children acquiring Kildin Saami as their first language. The youngest first language speakers belong to their parents' generation. However, there are a few young adult Kildin Saami who are trying to learn and use the language again. Language courses and other resources are not offered in great number, but are available nonetheless. In general, Kildin Saami individuals are interested in revitalising and maintaining the language and a few representatives take an active role in practical measures and relevant projects. For more detailed information on the current situation of Kildin Saami and ongoing revitalisation, see the descriptions in Rantala 1994, Sergeeva 1995, Sergeeva 2002, Scheller 2004, Scheller 2012 and further references mentioned there.

History of research and standardisation

The history of Kildin Saami standardisation is closely tied to the linguistic research history for the whole group of East-Saami languages. The following sections provide an overview focussing on contemporary written Kildin Saami. For a more detailed historical description of Kildin Saami literacy development, see Siegl et al. (in print). A comprehensive description of linguistic research history on Kildin Saami is in the works by the present author, and outlines are found in Sergeeva 2000 and Rantala 2005.

Orthography

Kildin Saami has been written since the end of the nineteenth century. However, the contemporary written language is not the result of a sustainable development of continuous modifications, because earlier standards were abolished in order to re-introduce completely new orthographies in the 1930s and 1970s. The history of Kildin Saami writing also includes a back and forth shift from Cyrillic to Latin to Cyrillic.

A translation of the Gospel of Matthew was the first book printed (partly) in Kildin Saami. It was written, with the help of native speaker consultants, in Cyrillic orthography by the Finnish linguist Arvid Genetz, in 1878. During the early Soviet period, a new, Latin-based orthography was successfully introduced. This new Kildin Saami written language was used in several textbooks for different subjects in the school education of Saami children[3] and in communist propaganda texts.[4] However, teaching Kildin Saami and the production of materials written in the language came to a complete standstill for decades when the official Soviet doctrine shifted away from minority language support in the late 1930s. The second new Cyrillic orthography for Kildin Saami was developed in the 1970s and 1980s by a working group of

Saami teachers and language activists led by the non-Saami educationalist and linguist Rimma Kuruch. This orthography (cf. *table,* p. 205) has since been used, in different variants, in dictionaries, in textbooks for elementary schools, as well as in several published literary texts for children and adults, and most recently even in the internet. An extensive presentation of Kildin Saami media history, including print and other written media, is included in Rießler (in print).

Documentation and description

Arvid Genetz was also the first researcher to work on an extensive linguistic documentation and description of Kildin Saami. Based on his Bible translation and a few other collected texts, he published a Kola Saami descriptive dictionary in 1891. Even the first short grammatical sketch of Kildin Saami (written in Hungarian and including comparative data of all Kola Saami languages) by Halász (1883) is based on Genetz' text collection. The first Kildin Saami grammar was written by Endukovskii (1937). The aim of this grammar was to support the training of Saami teachers. Consequently, the author applies a moderately prescriptive approach and uses the orthographic standard of that time, i.e. the Latin-based alphabet mentioned above.

T. I. Itkonen's (1958) comparative Kola Saami dialect dictionary (written in phonemic transcription and with translations into Finnish and German) is the most comprehensive source for Kildin Saami vocabulary. It comprises about 7,200 word stems, which are listed with all known cognate forms from different dialects of the four Kola Saami languages. A comprehensive descriptive grammar of Kildin Saami was produced by Kert in 1971. Other fragments of theoretical linguistic descriptions and several text collections (written in phonemic transcription and with translations into either Russian, Finnish or German) have also been published, mostly by researchers from Estonia, Finland, Hungary and Russia. Kildin Saami is thus not undocumented. Still, lexical and grammatical descriptions are incomplete and there is not much data available which reflect current language use.

Prescriptive dictionaries and grammars

There are three user's dictionaries of contemporary Kildin, published as Afanasyeva et al. 1985 (comprising almost 8,000 Kildin lemmas with Russian translations), Kert 1986 (a Kildin-Russian-Kildin school dictionary with about 3,000 lemmas in both directions)[5] and Sammallahti et al. 1991 (a Kildin-North Saami-Kildin dictionary with approximately 2,500 headwords in both directions).

The first two dictionaries mentioned above have more recently been combined and re-published as the *Electronic Saami dictionary* (Yur'ev 2003).[6]

The only user's grammar available for the modern Kildin Saami written standard is the grammatical sketch by Kuruch (1985). This prescriptive grammar contains a

short but comprehensive phonology, all main inflectional paradigms of nominals and verbs and a few very basic syntactic rules. The monograph on Kildin Saami orthographic rules by Kuruch et al. (1995) could in principle also be considered a prescriptive grammar because it includes detailed and almost complete descriptions of Kildin Saami phonology, morphophonology as well as (inflectional and derivational) morphology. The Sammallahti et al. (1991) dictionary includes a short appendix with basic inflectional paradigms.

The available printed dictionaries and grammars are clearly insufficient from both linguistic and pedagogical points of view. However, altogether they would still form a relatively solid base for teaching. The corpus of teaching materials even includes a few textbooks. Although those target elementary school pupils and the remaining few materials are practically useless for self-teaching by adults, interested Kildin Saami teachers have in fact a store of teaching materials at their disposal for preparing classes on different levels beyond elementary school level. Note that even today almost all active Kildin teachers are fluent L1-speakers with academic training. Today there is also funding available from different (Russian and non-Russian) sources for organising teaching at different places, including teacher compensation. According to my own observations, one important reason for the lack of progress in the quality of teaching is the generally low degree of basic pedagogical knowledge, combined with a lack of interest among teachers and the loss of attraction for learning the language among potential students, rather than the lack of contemporary teaching materials or funding (cf. also Scheller 2011: 101–102).

History of language planning

Language planning means the conscious interference in the development of a language in order to further advance its use in new domains (cf. Janich 2007). Consequently, Kildin Saami, a language which originally was only transmitted orally, was promoted as a written language, first in the 1930s and later in the 1970s, in order to become a functional communication device in Soviet society. The status of the language in society was supported by teaching it to children at school and using it in printed texts (status planning). To accomplish this task, a normalised written standard and new vocabulary adapted to modern society had to first be created (corpus planning).

The prescriptive dictionaries and grammars mentioned above are a product of the last fruitful period of language planning for Kildin Saami under the active participation of native speakers. In the 1970s, the language working group had already started producing preliminary teaching materials for Kildin Saami children. The aim was to create a linguistically and didactically well-founded written standard as the basis for a set of textbooks for different levels, as well as for dictionaries and didactic guides for Saami teachers. Although this ambitious goal was not reached completely (for different reasons), a considerable amount of materials has been produced, the most

important being the large Kildin Saami-Russian dictionary edited by Rimma Kuruch and co-workers (Afanasyeva et al. 1985), and a set of textbooks for grades 1 through 3, including accompanying didactic materials. Even the other Russian-Kildin-Russian school dictionary written by Georgii Kert (1986), the North Saami-Kildin Saami-North Saami wordlist written by Pekka Sammallahti and Anastasiia Khvorostukhina (1991) and other materials created recently by Kola Saami language activists (e. g. Sharshina et al. 2008) would not have been possible without the groundwork in language standardisation and orthography development by the Kildin Saami language working group.

Unfortunately, systematic work with language planning stopped in the late 1990s, when the Kildin Saami language working group at the Saami division of the Murmansk branch of the Russian Academy of Sciences disbanded. Reasons for terminating work included the lack of administrative and financial support after the collapse of the Soviet Union (Nina Afanasyeva, Rimma Kuruch p. c.), but perhaps also the ongoing tension between group collaborators and Saami language activists elsewhere. A sound description of the history of Kildin Saami language planning during the 1970s through the 1990s, including the founding, development and breakup of the Kildin Saami language working group, its achievements and its conflicts with outside language planners, has yet to be written. It can only be stated that true linguistic reasons, such as the prescription of dialectal forms not accepted by all speakers, are scarcely the trigger for the current conflicts. Unfortunately, all insider descriptions have so far been exclusively polemic, cf. a chapter in a book on the topic *The alphabet and orthography of the normative Kildin Saami language: History of the question* (translation MR) (Kuruch et al. 1995: 175–186) or personal statements provided to the present author in interviews or unpublished documents (e. g. by Aleksandra Antonova, Nina Afanasyeva, Ekaterina Korkina, Rimma Kuruch and Ekaterina Mechkina). Several scientific investigations of Kola Saami society also deal with this topic, but are rather biased in that they mainly portray the conflicts with the non-Saami researcher Rimma Kuruch, but disregard the fact that the history of the Kildin orthography has been quite successful overall (cf., e. g., discussions in Scheller 2004 and Øverland et al. 2012).

As a matter of fact, there is no professional infrastructure available today at a local level which could systematically support ongoing revitalisation attempts and continue the work with Kildin Saami language planning. Currently there are two local institutions evolving which could potentially fulfill this task. The first institution is a Saami research group called 'the Saami laboratory' led by the linguist Olga Ivanishcheva at the Murmansk State Humanities University.[7] Although Olga Ivanishcheva and her students conduct research on Kildin Saami, none of them has active proficiency in this (or any other Saami) language. So far they predominantly work on theoretical questions of Saami lexicology and terminology related to culture and society. It seems also unfavorable from a language planning perspective that the research group has neither Saami collaborators nor any close collaboration with Saami organisations.

The other institution, called the *Centre for Saami Competence*,[8] is itself a Saami institution located in Lovozero. It is provisionally led by the historian Maksim Kuchinskii and was officially registered as a nonprofit organisation in December 2011. According to project descriptions, practical language work will be but one target area of the centre's future work. Although the successful implementation of this aim hinges on the availability of permanent funding (which has not yet been provided) and the professional expertise of participating personnel, the centre should have the best requisites for also including a 'language centre', due to its close administrative and personal ties both to the local Saami community and to Saami representative organisations, as well as to Saami and non-Saami educational and academic research institutions in Russia and elsewhere.

A non-local research institution dealing partly with issues of Kildin Saami language planning is the *Centre for Sámi Language Technology* (Giellatekno)[9] at the University of Tromsø in Norway. Giellatekno provides, among other things, computer applications such as proofing tools, text processing tools, language learning programmes and digital dictionaries for Saami and other languages, and builds digital written language corpora for these languages.

Although North Sámi is the language that Giellatekno (the name actually means 'language technology' in North Saami) originally worked on and has most intensively researched, other Saami and even other northern languages have also always been the focus of work. Systematic work with Kildin Saami started in 1999 when Giellatekno launched a collaborative project between Trond Trosterud, Michael Everson and Rimma Kuruch to help provide the Kildin Saami language working group in Murmansk with the necessary digital infrastructure for using Kildin Saami on computers (cf. Everson 1999). Earlier, Everson and Trosterud had already proposed the addition of the Cyrillic Saami characters to the Unicode standard (ISO/IEC JTC1/ SC2/WG2 N1744 and N1590). Giellatekno also collaborated (together with Juhani Lehtiranta) on the creation of new fonts specifically for printing Kildin Saami at the Saami publishing house Davvi Girji in Kárášjohka[10] and (together with Everson) on the standardisation of character sets for Saami languages. Giellatekno's current work with Kildin Saami will be described below.

The Kildin Saami written language

The most fundamental achievement of Kildin Saami language planning has been the creation of a written standard that basically all Kildin Saami language users agree on today. A considerable corpus of printed texts using this standard is even available to future language planning activities, and this corpus is continuously growing.

The following section describes the standardisation of the Kildin Saami alphabet as an essential prerequisite for the creation of any digital infrastructure. Note that 'alphabet' refers to a set of graphic symbols (letters) representing sounds, whereas

'orthographic rules', which are not the topic of the present paper, are needed for combining the letters of an alphabet into words, word forms and larger text units in order to represent the written language in a conventionalised way.

After the introduction of the current Kildin Saami alphabet in the 1970s, it underwent several revisions reflecting improvements based on new and better insights into Kildin Saami phonology and morphophonology (for a description of the alphabet's history, including all revisions, see Kuruch et al. 1995: 175–186). The first, preliminary version of the alphabet was used in the Kildin Saami primer by Aleksandra Antonova (1982) and the dictionary by Georgii Kert (1986). The table *(p. 205)* presents the full set of characters after the last linguistically motivated revision leading to the version used in the comprehensive Kildin Saami-Russian dictionary (Afanasyeva et al. 1985).[11] The two letters in parentheses (14' and 18') are simply typographic replacements previous to the last orthography revision. This revision was meant as a compromise to settle a dispute over two alleged non-Cyrillic letters (Kuruch et al. 1995: 183). Although this most recent orthographic version has been used in the vast majority of published books up to today (i.e. in most textbooks published by the Kildin Saami language working group, in the North Saami-Kildin Saami-North Saami wordlist (Sammallahti et al. 1991), and in all Kildin Saami books published by Davvi Girji in Norway during the 1990s), the two replaced letters were later abandoned again and this alphabet variant seems to be completely out of use today. A third orthographic variant, used by several current authors, does not include either letters 14, 18, 14' or 18' and is similar to the initial variant introduced in print by Antonova (1982).

Shortcomings, gaps and open questions in the standardised alphabet

The Kildin Saami alphabet is the result of the continuous groundwork completed by the Kildin Saami language working group since the 1970s. Saami teachers and activists were included in this work under the leadership of the non-Saami academic Rimma Kuruch (see also the section on language planning above). The successful creation and establishment of a written norm is a major achievement in the history of Kildin Saami revitalisation. It is the most fundamental basis for all recent, current and future work with language planning and language technology. The following discussion of possible shortcomings, gaps and questions does not intend to diminish the efforts by Kuruch and her collaborators. On the contrary, the new creation of an alphabet is far from trivial and in most cases its practical implementation will need a sustained period of evaluation and possible revisions.

Linguistic shortcomings. From a linguistic perspective, this alphabet with a number of modified letters, like the vowels with macron (marking length, e.g. ӭ) and diaeresis (marking so-called 'half-palatalisation', e. g. ӓ) or the consonants with tail (marking devoicing, e.g. ӈ) and hook (the velar nasal ӈ) seems unnecessarily complicated because some marginal oppositions in Kildin Saami phonology are overempha-

sised, e.g. the length of vowels and the voicelessness of sonorant consonants. On the other hand, some rather relevant phonological distinctions are represented unsystematically. This is especially true for the difference between palatalisation as a secondary modification of consonants and the primary place of palatal articulation of other consonants. This important phonological distinction is completely blurred by the unsystematic use of the palatalisation and half-palatalisation marks or di- and triglyphs for representing the respective sounds (cf. Kuz'menko et al. 2012). It is most likely that the alphabet creators' original intention was to stress the linguistic differences between Russian and Kildin Saami, as they are perceived by native Saami speakers, who are also fully literate in the majority language Russian. Incomplete and even wrong linguistic analyses from that time also seemed to have played a role. For Saami members among the working group, who were all well-educated and fluent bilinguals, the alphabet was surely a useful tool for representing their native language in written form. However, teaching Kildin Saami pronunciation as a foreign language with the help of this alphabet is clearly a didactic challenge. The introduction of a typographically complicated alphabet, in combination with missing linguistic advice and the teachers' generally poor knowledge of multilingual education, has likely caused severe difficulties in Kildin Saami language education from the beginning until today.

Typographic shortcomings. A high level of creativity in the construction of new letter symbols was generally typical of minority language planners in the Soviet Union (Trosterud 1996). [12] From the point of view of standardisation, the large amount of special letters in the new Cyrillic alphabet for Kildin is definitely a challenge, both for language learners and language technologists. The Kildin Saami alphabet (main variant, see *table* on p. 205) contains eight letter bases which are not included in the Russian alphabet. Adding a macron on the vowels adds another nine letters, plus two more letters with diaeresis. Altogether, the additional 19 special Kildin Saami letter symbols increase the basic modern Russian alphabet's 33 letters by more than 50 %.

The high number of Kildin Saami special letters and their typography has already resulted in technical shortcomings affecting work with written Kildin Saami in the pre-digital age. The resulting problems concern mainly the input of letters with different technical devices but sometimes even their typographic appearance on the output side.

The special Kildin Saami letters were created according to different principles. In most cases, a Russian letter was typographically modified. Thus, for instance, adding a diacritic MACRON (ˉ) as a marker of vowel length was clearly borrowed from a similar rule used in phonemic scripts. Several other modified letters also borrowed their meaning from other known alphabets, for instance the CYRILLIC LETTER EN WITH HOOK (Ӈ ӈ), which marks a voiced velar nasal in different Uralic (and several other) languages. The CYRILLIC LETTER SHHA (Һ һ) and the CYRILLIC LETTER JE (Ј ј) were also borrowed from existing alphabets of other languages. The former is used in the Cyrillic alphabets of several Turkic languages, where it also marks a

voiceless fricative sound. The latter comes from the Cyrillic alphabets of Serbian and Macedonian, where it marks an entirely different sound. Note that in current non-professional arguments against these two letters, one of the main reasons for not using them is their alleged status as 'Latin letters'. This is obviously not entirely true, even though they are certainly better known from the Latin alphabet.

A few other letters are inventions rather than borrowings, such as the CYRILLIC LETTER SEMISOFT SIGN (Ҍ ҍ) for the (phonologically misinterpreted) so-called half-palatalisation. This letter resembles the Old Slavonic CYRILLIC LETTER YAT (Ѣ ѣ), but is typographically slightly different from it.[13] The use of diacritic DIAERESIS (¨) for the (phonologically mis-interpreted) so-called half-palatal vowels CYRILLIC LETTER A WITH DIAERESIS (Ӓ ӓ) and CYRILLIC LETTER E WITH DIAERESIS (Ӭ ӭ) also follows a principle which the creators of the alphabet had developed themselves.

₁ А а	₂ А̄ а̄	₃ Ӓ ӓ	₄ Б б	₅ В в	₆ Г г	₇ Д д
0410 0430		04D2 04D3	0411 0431	0412 0432	0413 0433	0414 0434
₈ Е е	₉ Е̄ е̄	₁₀ Ё ё	₁₁ Ё̄ ё̄	₁₂ Ж ж	₁₃ З з	₁₄ һ һ
0415 0435		0401 0451		0416 0436	0417 0437	04BA O4BB
₁₄'(')	₁₅ И и	₁₆ Ӣ ӣ	₁₇ Й й	₁₈ Ј ј	₁₈'(Ӥ, ӥ)	₁₉ К к
02BC	0418 0438	04E2 04E3	0419 0439	0408 0458	048A 048B	041A 043A
₂₀ Л л	₂₁ Ӆ ӆ	₂₂ М м	₂₃ Ӎ ӎ	₂₄ Н н	₂₅ Ӊ ӊ	₂₆ Ӈ ӈ
041B 043B	04C5 04C6	041C 043C	04CD 04CE	041D 043D	04C9 04CA	04C7 04C8
₂₇ О о	₂₈ О̄ о̄	₂₉ П п	₃₀ Р р	₃₁ Ҏ ҏ	₃₂ С с	₃₃ Т т
041E 043E		041F 043F	0420 0440	048E 048F	0421 0441	0422 0442
₃₄ У у	₃₅ Ӯ ӯ	₃₆ Ф ф	₃₇ Х х	₃₈ Ц ц	₃₉ Ч ч	₄₀ Ш ш
0423 0443	04EE 04EF	0424 0444	0425 0445	0426 0446	0427 0447	0428 0448
₄₁ Щ щ	₄₂ Ъ ъ	₄₃ Ы ы	₄₄ Ь ь	₄₅ Ҍ ҍ	₄₆ Э э	₄₇ Э̄ э̄
0429 0449	042A 044A	042B 044B	042C 044C	048C 048D	042D 044D	
₄₈ Ӭ ӭ	₄₉ Ю ю	₅₀ Ю̄ ю̄	₅₁ Я я	₅₂ Я̄ я̄		
04EC 04ED	042E 044E		042F 044F			

The current Kildin Saami alphabet, including capital and lowercase letters.
The subscript number to the left of each letter pair is used for reference elsewhere in this paper.
The Unicode (Hex) is given below each letter. If there is no Unicode, the letter has to be composed either with the COMBINING DIAERESIS (Unicode Hex 00A8), or with the COMBINING MACRON (Unicode Hex 00AF).

The input of the large amount of special letters not available on commonly used digital writing devices continues to be the main challenge. Earlier, special letters or letter modifiers (tail, hook, etc.) not available on the Russian typewriter were often added manually afterwards in the manuscript. This solution is still sometimes used in combination with a computer (e. g. in the second issue of the journal *Sām̄* 'Saami', published as Samizdat by Korkina and Galkina in 2005). For manuscript creation by the Kildin Saami language working group, a tailor-made typewriter including all

needed letters was produced (Aleksandra Antonova, Nina Afanasyeva p. c.). In the computer age, especially after the breakthrough of the international standard Unicode (cf. Allen 2011) in practically all commonly used applications (office programmes, internet browsers) this coding problem is solved, in principle. However, inputting the special Cyrillic letters of the Kildin Saami alphabet, combined with incomplete or typographically non-appealing fonts on the output side, remain the most crucial problems.

Regardless of all linguistic and typographic shortcomings in the large inventory of Kildin Saami special letters, I believe that any attempt to revise the orthography 'from above' would create even more confusion and perhaps conflicts. Computational problems with Unicode, font typography and input devices will likely be solved soon. The teaching of complicated (or even linguistically unsystematic) writing systems is in principle possible, as several cases from other parts of the world show. Note also that Saami teachers and other language users seem to identify themselves with the alphabet specifically developed for their native language. The outside researcher working with language planning should thus concentrate on the provision of the necessary basis for revitalisation (such as adequate linguistic description and language technology), rather than interferring in orthography revisions.

Sorting order. Although the typographic shape of Kildin Saami letters has been standardised and all letters and diacritics are included in Unicode, a standardised sorting order for all letters was never suggested explicitly by the authors of the alphabet.

The standardised letter inventory of the Kildin Saami alphabet can be found in different textbooks for schools (e. g. in Antonova 1990: 130–131), in the three dictionaries (Afanasyeva et al. 1985: 16; Kert 1986: 9; Sammallahti et al. 1991: 6), as well as in the monograph on Kildin Saami orthographic rules (Kuruch et al. 1995: 7). Deducing from the alphabets presented there, authors agree on a general sorting order similar to the modern Russian alphabet, with the Kildin modified letters following the respective basic letters. However, although the two vowel letters with diaeresis are included, following the non-modified vowel letters, vowel letters with macron are not listed. Instead the macron is explained separately as a marker of vowel length.

It would in principle be possible to exclude the vowel letters with a macron from the alphabet, i. e., to not handle them as separate letters.[14] A standardised sorting order is nevertheless needed, not only for lemma lists in dictionaries, but also for sorting routines on the computer.

In the table *(p. 205)*, I propose a sorting order in which all vowel letters with a macron directly follow the respective plain vowel letters but precede letters with diaeresis. A very similar ordering of letters with macrons and diaereses was also suggested by Everson (2002).[15]

Infrastructure for endangered languages in the digital age

Written language production and reception in the computer age has become unthinkable without the internet. Therefore the survival of written languages hinges on the availability of the relevant digital infrastructure. However, computer technology and the internet is not exclusively a threshold that a language must exceed to survive modernisation. On the contrary, these technologies also offer valuable resources for vitalising active production and perception and thus reversing language shift. The internet (including e-mail, social networks, etc.) is an excellent long-distance communication device and is constantly becoming cheaper and more widespread. The internet's role as a local information channel is growing fast as it is becoming more easily available to more people. Last but not least, the internet makes text production easier as it lowers publication costs. One and a half decades ago (long before Wikipedia and Facebook where launched), Trosterud (1996) explicitly described these important characteristics of the future internet as a crucial resource for vitalising the endangered Northern languages. Trosterud was perhaps too optimistic about how fast this development would take place. However, the developments during the last 15 years bear out his vision.

Basically, 'general digital infrastructure' means technology that makes it possible to use a language in today's computerised society, especially in the governmental context, in education, or in modern publishing activities. Any general digital infrastructure must thus include: (1) computer fonts covering all letters of the specific language in question, (2) keyboards for input of these letters on the common computer platforms, (3) annotated language corpora as data pools for language technology, (4) multi-language dictionaries in electronic form, and (5) proofing tools for the most important office programmes.

In the Soviet context, many written minority languages were traditionally used in schools and partly also in cultural life. However, today most minority languages of the former Soviet Union have practically no access to language technology resources (with very few exceptions). Without a digital infrastructure in place, these languages will not make the necessary transition to being functional in modern information and communication technology, and will thus fall out of use in public life. In the same vein, any official statements about recognising and supporting a minority language are only empty declarations as long as we are not able to use computers to write letters in the language, or to correct text or find the correct terminology. Until this happens, the language cannot be put into use in the administration of modern societies.

Available digital infrastructure for languages can be divided hierarchically into three main levels, each building on the previous one. On the first level is the generation of languages which have keyboard layouts, fonts, sorting routines, as well as time and date formats already available. Kildin Saami is only in a preliminary stage of becoming a member of this language generation.

Spellcheckers, automatic hyphenation and electronic dictionaries characterise languages of the second generation, to which North Saami already belongs. South and Lule Saami have also recently made the jump to this generation and have these tools available now.

Building on that, electronic thesauri, machine translation and speech technology can be created for third generation languages, such as Russian, Swedish, Finnish and Estonian.

To make the work on the mentioned infrastructure and tools a sustained base for further development, it is most crucial that the applications created are open source, without any third party owning even part of the source and able to block similar efforts in the future. That means that all gathered knowledge should be made explicit by providing an open project documentation. Furthermore, knowledge inside the relevant language communities has to be fostered. A permanent and sustained infrastructure as defined here is important in preventing existing results from being wasted in the long term. It should be possible for everyone, with reasonable effort, to pick up the thread again at any later point and continue to build on the work already done.

Computer literacy. Creating and providing digital infrastructure for an endangered language is useful for revitalisation only if the targeted users, i.e. members of the endangered language community, can access and use it. Purely technical restrictions, like access to computers and the internet, will hopefully be irrelevant in the near future of Kildin Saami. Although computerisation in rural communities of the Russian North is slower than optimistic forecasts earlier suggested, there is still clearly continuous progress. A much more serious problem – and one rarely addressed in connection with revitalisation and the pedagogy of endangered languages – is the degree of computer (and general digital) literacy among language users.

Computer skills are a logical precondition for using any kind of digital linguistic infrastructure. Since most Kildin Saami individuals who can read and write their native language use computers (at least occasionally) nowadays, this precondition is generally met. In the current Kildin Saami context,[16] however, computer literacy is restricted to a very basic level and computer skills are almost exclusively restricted to the use of very specific applications and a few very well-defined simple tasks. Most typically, these include writing and formatting shorter text documents in Microsoft Word for Windows, browsing the internet – including web-based social networks and e-mailboxes – with Microsoft Internet Explorer as well as using a few multimedia applications included in Windows. Needless to say, any additional knowledge and ability beyond basic skills would constitute a significant asset in using digital resources more efficiently and hence more successfully for revitalisation.

Note that the problem of computer illiteracy is far from being specific to Kildin Saami or other endangered language communities. On the contrary, feeling uncomfortable with using anything other than a few familiar applications for a restricted set of specific tasks seems rather common, even for many people using computers professionally every day. However, working less efficiently as a linguist, for instance,

by formatting a research paper in an office programme manually from paragraph to paragraph or inserting and formatting each reference individually and by hand seems trivial compared to computer illiteracy in the endangered language context, where ignoring the very few available resources for revitalisation is much more crucial.

It is definitely the case that the creator and provider of any sustainable digital linguistic infrastructure should also put effort into promoting computer literacy among native language users. For Kildin Saami language activists and potential collaborators in future language technology projects, training has been provided occasionally both by 'Kola Sami Documentation Project' *(*KSDP) and Giellatekno. Participants had very different levels of computer literacy, ranging from using the computer for the first time to being well-acquainted with office programmes and internet-based communication such as e-mail, instant message programmes and social networks. Teaching thus ranged from a general introduction to working with computers (the difference between an operating system and an application running on that system, installing and using common applications for language work) to introduction to Unicode and the proper use of character encoding in applications and fonts, as well as the use of internet-based communication software (Skype, Google, etc.) for file sharing and long-distance collaborative work. Courses even dealt with general database structure and text technology applied for instance in wiki-markup language or XML-based lexicography.

Documentation and technology. Documentary linguistics is a relatively new and evolving field primarily concerned with language documentation as a comprehensive, multi-faceted and multi-purpose record of linguistic practices characteristic of the investigated speech community (cf. Himmelmann 2006: 1 and other contributions in Gippert et al. 2006). Although documentary linguistics evolved from traditional field linguistic methodology with the primary aim of providing more and better data on the world's linguistic diversity for future research in theoretical linguistics, the field is currently developing into a linguistic sub-discipline of its own. Among the most important purposes of language documentation is the provision of data for further research on and for endangered languages, for both further theoretical and applied research, as well as for direct use by the investigated language community. Ideally, the data pool provided by the documentary linguist includes a comprehensive, deeply annotated and easily accessible multimedia corpus of a spoken language. Recent calls for project grant applications, e.g. by DoBeS[17] or HRELP,[18] and numerous recent conferences or workshops organised by a variety of institutions dealing with documentary linguistics, also explicitly include technical, methodological and theoretical issues connected to endangered language revival. If this collaborative approach to documentary linguistics is to be more than just window-dressing, language documentation must first of all also aim at working together with and obtaining useful results for the investigated community.

Beside methodologies and techniques related to archiving and corpus building, which belong to applied research per se, documentary linguistics interfaces with

different aspects of language planning and language technology for endangered languages. Here, language technology is understood as the functional application of computational science as it is aimed at analysing and generating natural language in various ways and for a variety of purposes. Machine-based translation or automatic language analysers are but two examples of such practical applications. Whereas the documentary linguist provides language corpora and linguistic analyses necessary for the computational modeling of the language in question, the language technologist applies formal-descriptive linguistic and corpus linguistic methods to the programming of machine readable morphological, syntactic and lexical descriptions of the relevant languages, and thereby creates tools for effectively analysing language corpora and carrying out better linguistic documentation and description.

Although all combined efforts between language technology and documentary linguistics are also directly profitable in the revitalisation of endangered languages, these two fields rarely meet in practice today. One possible explanation is that documentary linguistics predominantly deals with the documentation of spoken language, whereas applications created in language technology are almost exclusively based on written language corpora. Another point is that mainstream approaches in documentary linguistics still predominantly focus on the improvement of methodologies and techniques for field- or corpus based empirical data analyses of endangered languages, rather than specifically including applied research towards providing data and revitalisation.

The theoretical bias in documentary linguistics, which is likely a relic of the typologically oriented theoretical-functional linguistic school from which most current documentary linguists descend, is well-illustrated in the overemphasised need for thorough morphosyntactic analyses in corpus annotation (frequently carried out by researchers manually). While such annotations can be quite precise, they require a lot of time and leave much room for inconsistencies and mistakes. Mostly, a basic transcription, a basic translation into the main lingua franca and a basic grammatical description accompanying the corpus seem sufficient in order to make the language documentation useful, i.e. readable and accessible to further linguistic analyses or practical applications. This is especially true for languages like Kildin Saami, which already have an established orthography and a basic linguistic description available. Rather than providing our corpus with phonological and morphosyntactic analyses and English translations, like many other similar documentation projects would do, we believe that transcriptions using the established Cyrillic writing system and providing Russian translations is entirely adequate from the perspective of making our documentation useful for the broadest possible future applications, including language technology, while also utilising the limited time and financial resources we have at our disposal for maximum effect.

Building a digital infrastructure for Kildin Saami

Regarding digital infrastructure, Kildin Saami is still at the starting block. However, a comprehensive annotated multimedia corpus is in the works by KSDP. Fonts and preliminary keyboard layouts are already available and pilot work on digital corpora, electronic dictionaries and even grammatical analysers has been done. The step to a sustained second-generation infrastructure (similar to what is currently being done for Lule and South Sami, as well as Komi-Zyrian) is the short-term objective of KSDP and Giellatekno. Later, even more advanced tools will hopefully be created by us or other projects.

The remaining sections briefly describe work in progress on the creation of an improved digital infrastructure for Kildin Saami, in which the present author is involved.

Kildin Saami keyboards. Work on functional and effective keyboards for Kildin Saami is currently being carried out as part of the 'Barents keyboard project' (which is itself part of a pilot project on 'Digital infrastructure and language technology for indigenous languages of Northwest Russia').[19] Although there are keyboard solutions available which make it possible in principle to produce text in Kildin Saami on the computer, the praxis shows that better keyboard drivers for the most common platforms are needed in order to further support active use of written Kildin Saami on the computer.

The keyboards currently developed build on standardised designs based on the existing Russian keyboards. They need to be user-friendly during installation on all common platforms (Windows, Mac, Linux) and will be equipped with a user guide in Russian. Furthermore, it is important to provide two variants: one variant uses Russian as the main language and has all special Kildin Saami letters available only through key combinations. The other keyboard, focussing on the minority language, has the most frequent Kildin Saami letters visible on top (replacing non-frequent Russian letters).

This work is planned to be finished in summer 2012. Besides publishing keyboard drivers for free downloading, the project will also propose official keyboard standards to be submitted to the Russian standardisation body. The envisioned goal is to introduce the keyboards into the most important operational systems, like the keyboard drivers for Saami languages in the Nordic countries, which are included in Windows, Mac and Linux.

Note, however that our keyboard solutions can only solve the problem of character input with the correct Unicode. Serious problems with the representation of Kildin Saami on the computer remain, but are out of our hands to deal with. Many common fonts still miss several (or all) of the Kildin Saami special glyphs, and show squares instead. Another problem is that many applications do not combine diacritics with the respective letters.

Kildin Saami digital language corpora. Although many hours of recorded Kildin Saami spoken language are stored at different archives, mainly in Estonia, Finland, and Russia, the first linguistic documentation focussing on digital data was the Kola Saami Documentation Project, a DoBeS project led by Jurij Kusmenko and the present author (cf. Rießler et al. 2007).[20] The project created a multimedia documentation of Kildin and the other Kola Saami languages.[21] Although the regular funding period for the project ended in 2010, the present author continues to work on archive building.

Currently, most archived text recordings are only provided with cataloguing metadata and preliminary orthographic transcriptions. In its current state, the corpus can only be further annotated, presented and searched using the tools created and implemented by the Language Archiving Technology team at DoBeS/MPI Nijmegen.[22] These tools have been created specifically to meet the needs of quite divergent linguistic and anthropological documentation projects. In this, it is crucial, and in accordance with current methods in documentary linguistics, that annotations are linked to multimedia, discoverable via rich metadata and accessible via user-friendly interfaces.

One significant difference between KSDP and common practice in other documentation projects is that we use orthographic representations of spoken language data rather than phonological transcriptions. It is our plan to benefit from methodologies and computational linguistic infrastructure and tools created by Giellatekno specifically for the preparation of a searchable corpus by automatised annotations. The spoken language corpus will later be able to be joined with corpora of written texts. This will result in a significantly higher volume of texts for improved corpus research in the future. Finally, the resulting larger corpus infrastructure will later also be able to provide better resources for future (theoretical and applied) research. For example, lexical resources, such as word form frequency lists, letter frequency lists, and part-of-speech frequency lists will be created automatically. The corpus can also be lemmatised and parsed automatically and thus be used for improving any future lexicographic work.

Last but not least, the multimedia corpus, including acoustic and video data, might in the future also provide a basis for contemporary teaching materials.

Kildin Saami dictionary project. The basis of Giellatekno's digital dictionary infrastructure for Kildin Saami[23] is identical to the Kildin Saami-Russian part of Yur'ev's (2003) *Electronic Saami Dictionary* (see also above), which the author kindly shared with us in digital form. The digital dictionary shows some drawbacks in regard to content, usability and availability. Among other things, the dictionary is not distributed freely, it can be installed only on Windows systems with a Russian language localisation, and its search function needs a special font due to the incompatibility of the dictionary with Unicode. Therefore, we decided to focus on transforming the existing data into a lexicographically better structured XML database, which will later allow the application of XSLT-scripts to export into different kinds of platform-independent dictionary or other teaching applications, for which Giellatekno already has the necessary infrastructure available, see, for instance, the LEKSA programme

for teaching vocabulary[24] or the online multi-lingual dictionary portal for a variety of languages.[25] It will also be possible to extend the database with example phrases from the automatically analysed text corpora (see above).

Yur'ev (2003) consists of lexical data found in the two printed Kildin-Russian dictionaries and translations in both directions. However, while data from both sources (i. e. Afanasyeva et al. 1985 and Kert 1986) are included in the Russian translations, the reverse translations are restricted to the Russian-Kildin Saami part of Kert 1986. We plan to complete the Russian-Kildin translations by converting the much larger Kildin-Russian part of the dictionary. In fact, a comprehensive dictionary translating from Russian into Kildin Saami or explanations in Kildin Saami is needed most urgently by students or other language users who are not fluent first-language speakers for creating texts. Needless to say, merely reversing a dictionary will not automatically result in an adequate dictionary, unless the lexicographer has put a great deal of manual effort into refining the resulting database (cf. also Trosterud 2001).

Conclusion and prospects

"Rather than being written for the average American web surfer, the future web page will be written for well-known customers, fellow citizens, or for neighbours. Thus, future web pages will reflect human life on this planet in all its colourful varieties" (Trosterud 1996).

The wish, manifested in this quote from only fifteen years ago, to reflect our world's linguistic variation in the digital world, has already come true if one thinks about the many threads posted on digital social networks every day in Saami and other minority languages, or about the globally free and collaborative internet encyclopedia project hosting Wikipedias, not only for the 'average American web surfer', but also for speakers of much smaller national languages and even non-national minority languages like North Saami.

Quite recently, a Kildin Saami test Wikipedia was launched. Let us make it possible for the Kildin Saami language to become fully functioning as a language of education, fine arts and communication in our digitally globalised world, following the path of other formally alleged 'minor languages' like North Saami, Estonian and Finnish, and finally disprove Nervander's disbelief in the 'capabilities' of small languages.

Notes

1   Kristina Kotcheva, Beverley Stewart, Nick Thieberger, Joshua Wilbur and the editors of the present volume provided valuable comments and suggestions on earlier drafts of this chapter. Special thanks are due to Sjur Moshagen and Trond Trosterud for discussing specific issues on digital infrastructure for Saami languages. Needless to say all remaining errors are my own.

2   Other common spellings are Sámi or Sami.

3   E. g. Popova 1934: *Olghš škola varas arifmetika opnuvǝm kniga* 'Textbook in arithmetics for primary school'.

4   E. g. Valerštejn 1934: *Mi lij mogka industrializacija jemnest* 'Industrialisation of the country – what does this mean'.

5   But note that the subtitle on page 1 states '4000 words', which is not exactly true.

6   This dictionary, together with a keyboard driver and the font RuSaDic (both non-Unicode), are distributed by the author, on CD-ROM, or via the website http://saami.su/saami/rusadic.html.

7   http://www.mshu.edu.ru/

8   http://saami-tied.ru/

9   http://giellatekno.uit.no/

10 http://www.davvi.no/

11 In the Kildin Saami community and for some researchers, this dictionary is also often referred to as *Kuruch's dictionary* – after the leader of the working group and main editor (Russian: *redaktor*) – or as the *Red dictionary* – after the hard cover version's prominent red colour (perhaps also with an allusion to the included extensive Soviet terminology created for the first time for Kildin Saami?). According to common standards the authors are listed here alphabetically, just as in the imprint on page 4 of the dictionary.

12 A similar situation has led to the same kind of consequences for today's language learners and language technologists concerning the Skolt Saami alphabet developed by linguists in Finland in the 1970s.

13 According to Aleksandra Antonova (p. c.) this difference was intentional.

14 Similarly, letters with (optional) accent diacritics in Norwegian, like in *òg* 'also' and *vær* 'weather' are never listed as separate letters of the alphabet.

15 Note that the order of a few other letters in Everson's document is different from the other alphabets mentioned and published in Afanasyeva et al. 1985; Kert 1986; Sammallahti et al. 1991; Antonova 1990 and Kuruch et al. 1995.

16 This is according to my own observation. To my knowledge, there is no statistical data available, as surveys on computer literacy in connection to general language use and revitalisation are scarcely ever included in language sociological research.

17 http://www.mpi.nl/dobes/

18 http://www.hrelp.org/

19 Project leaders are Sjur Moshagen (Norwegian Saami Parliament) and Trond Trosterud (Giellatekno), see http://giellatekno.uit.no/plan/barents/doc/barentspresentation.eng.html and the project documentation at http://victorio.uit.no/cgi-bin/wiki/index.php/Barents_keyboard_project/.

20 Other collaborators contributing to this language documentation were Anna Afanasyeva, Anja Behnke, Svetlana Danilova, Andrei Dubovtsev, Elena Karvovskaya, Kristina Kotcheva, Elisabeth Scheller, Nina Sharshina, Ganna Vinogradova, Joshua Wilbur, Evgeniya Zhivotova and Nadezhda Zolotuchina.

21 http://corpus1.mpi.nl/ds/imdi_browser?openpath=MPI363060%23

22 http://tla.mpi.nl/tools/tla-tools/

23 The project is carried out in collaboration between Ciprian Gerstenberger (Giellatekno) and the present author.

24 LEKSA is only one part of the teaching platform Oahpa!. For the North Saami version, see http://oahpa.no/davvi/

25 http://victorio.uit.no/webdict

References

Afanasyeva, Nina E., Rimma D. Kuruch, Ekaterina I. Mechkina, Aleksandra A. Antonova, Lazer' D. Jakovlev and Boris A. Gluchov 1985. *Saamsko-russkii slovar' = Sām'-rūšš soagknehk'.*[Saami-Russian dictionary]. Moskva: Russkii yazyk.

Aikio, Ante 2003. The Geographical and Sociolinguistic Situation. In *Siiddastallan: from Lapp communities to modern Sámi life.* J. Pennanen and K. Näkkäläjärvi (eds.), 34–40. Publication of the Inari Sámi Museum 5. Jyväskylä: Siida Sámi Museum.

Allen, Julie D. 2011. *The Unicode Standard, version 6.0.* The Unicode Consortium, Mountain View. http://www.unicode.org/versions/Unicode6.0.0/.

Antonova, Aleksandra A. 1982. *Sām bukvar': dlia 1 klassa saamskikh shkol.* [Saami ABC book for the 1st class of Saami schools]. Leningrad: Prosveshchenie.

— 1990. *Sām' bukvar': dlia 1 klassa saamskikh shkol.* 2e izdanie, dorabotannoe. [Saami ABC book for the 1st class of Saami schools, second revised edition]. Leningrad: Prosveshchenie.

Endiukovskii, Aleksandr G. 1937. Saamskii (loparskii) yazyk [The Sami (Lappish) language]. In *Yazyki i pis'mennost' narodov Severa I: yazyki i pis'mennost' samoedskikh i finno-ugorskikh narodov.* [Language and writing system of the Samojed and Finno-ugric peoples]. G. N. Prokof'ev (ed.), 125–162. Trudy po lingvistike. Nauchno-Issledovatel'skaia Associaciia Instituta Narodov Severa imeni P. G. Smidovicha. Moskva: Gosudarstvennoe uchebnopedagogicheskoe izdatel'stvo.

Everson, Michael 1999. *When is a Descender not a Descender: Kildin Sámi Voiceless Consonants.* Expert Contribution.
http://www.hum.uit.no/a/trond/kildinbackgr.pdf

— 2002. The Alphabets of Europe: Kildin Sami. Manuscript.
http://www.evertype.com/alphabets/kildin-sami.pdf

Genetz, Arvid 1878. *Mah'tveest Pas'-Evangeli: Samas = Evangelie ot Matfeja (na russko-loparskom yazyk)* [The Gospel of Matheus (in the Russian-Lappish language)]. Izdannoe Obshchestvom Rasprostraneniia Biblii v Veliko-Britanii i v drugikh stranach [edited by the British Society for distribution of the Bible in other countries]. Translation A. Genetz. Helsinki: British and Foreign Bible Society.

— 1891. *Kuollan lapin murteiden sanakirja ynnä kielennäytteitä = Wörterbuch der Kola-Lappischen Dialekte nebst Sprachproben.* Bidrag till kännedom af Finlands natur och folk 50. Helsinki: Finska Vetenskaps-Societeten.

Gippert, Jost, Ulrike Mosel and Nikolaus Himmelmann (eds.) 2006. *Essentials of Language Documentation.* Trends in Linguistics. Studies and Monographs 178. Berlin: Mouton de Gruyter.

Gröndahl, Satu 2009. Eposet som nationalsymbol bland finsk-ugriska folk [The epic as national symbol among Finno-ugric peoples]. *Annales Academiæ Regiæ Scientiarum Upsaliensis* 2007–2008, 37: 55–72.

Halász, Ignácz 1883. Orosz-lapp nyelvtani vázlat [An outline of Russian-Lappish grammar]. *Nyelvtudományi közlemények* 17: 1–45.

Himmelmann, Nikolaus P. 2006. Language Documentation: What is it and what is it good for? In *Essentials of Language Documentation*. Trends in Linguistics. Studies and Monographs 178. J. Gippert, U. Mosel and N. Himmelmann (eds.), 1–30. Berlin: Mouton de Gruyter.

Itkonen, Toivo Immanuel 1958. *Koltan-ja kuolanlapin sanakirja = Wörterbuch des Kolta- und Kolalappischen*. Vol. 15. Lexica Societatis Fenno-Ugricae. Helsinki: Suomalais-Ugrilainen Seura.

Janich, Nina 2007. Sprachplanung. In *Angewandte Linguistik: ein Lehrbuch*. K. Knapp (ed.), 481–501. Tübingen: Francke.

Kert, Georgii M. 1971. *Saamskii yazyk (kil'dinskii dialekt): fonetika, morfologiia, sintaksis*.[The Saami language (kildin dialect), phonetics, morphology, syntax]. Leningrad: Nauka.

— 1986. *Slovar' saamsko-russkii i russko-saamskii*. [Saami-Russian and Russian-Saami dictionary]. Leningrad: Prosveshchenie.

Kuruch, Rimma D. 1985. Kratkii grammaticheskii ocherk saamskogo yazyka. [Short grammatical summary of the Saami language]. In *Saamsko-russkii slovar' = Sām'-rūšš soagknehk'* [Saami-russian dictionary]. N. E. Afanas'eva et al. (eds.), 529–567. Moskva: Russkii yazyk.

Kuruch, Rimma D., Nina E. Afanas'eva and Iraida V. Vinogradova 1995. *Pravila orfografii i punktuatsii saamskogo yazyka* [Orthographic rules of the Saami language]. Murmansk: Murmanskii sektor lingvisticheskikh problem finno-ugorskikh narodnostei Krainego Severa Instituta Yazykoznaniia Rossiisskoi Akademii nauk.

Kuz'menko, Yurii [Jurij Kusmenko] and Mikhael' Rissler [Michael Rießler] 2012. K voprosy o tverdykh, miakhkikh i polumiakhkikh soglasnykh v kol'skosaamskom [On the problem of hard, soft and half-soft consonants in the Kola Saami language]. In *Acta Fenno-Lapponica Petropolitana*. Vol. 8: 1. N. V. Kuznetsova, V. S. Luleshov and M. Z. Muslimov (eds.), 20–41. Trudy Instituta lingvisticheskikh issledovanii RAN [Proceedings of the Institute of linguistic research of the Russian Academy of Sciences]. Sankt Peterburg: Nauka.

Øverland, Indra and Mikkel Berg-Nordlie 2012. *Bridging the Divides: Ethno-Political Leadership among the Russian Sámi*. Oxford: Berghahn Books.

Rantala, Leif 1994. Samerna på Kolahalvön: deras situation i dag. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 85: 200–204.

— 2005. Saami studies: Russia/USSR. In *The Saami: a Cultural Encyclopaedia*. U.-M. Kulonen et al. (ed.), 363–365. Suomalaisen Kirjallisuuden Seuran toimituksia 925. Helsinki: Suomalaisen Kirjallisuuden Seura.

Rantala, Leif and Aleftina Sergina 2009. *Áhkkila sápmelaččat: Oanehis muitalus sámejoavkku birra, man maŋimuš sámegielalaš olmmoš jámii 29.12.2003* [Akkala

Saami: a short history about a Saami group, whose last speaker died 29.12.2003]. Roavvenjárga.

Rießler, Michael (in print). Le same de Kildin face aux médias, une revitalisation de la langue? In *L'Image du Sápmi 2*. K. Andersson (ed.). Humanistica Oerebroensia. Artes et linguae. Örebro: Örebro universitet.

Rießler, Michael and Joshua Wilbur 2007. Documenting the Endangered Kola Saami languages. In *Språk og språkforhold i Sápmi*. T. Bull, J. Kusmenko and M. Rießler (eds.), 39–82. Berliner Beiträge zur Skandinavistik 11. Berlin: Humboldt-Universität zu Berlin.

Sammallahti, Pekka 1998. *The Saami Languages: an Introduction*. Kárášjohka: Davvi girji.

Sammallahti, Pekka and Anastasiia Khvorostukhina 1991. *Unna sámi-saam' saam'-sámi sátnegirjjáš = Udc' sámi-saam' saam'-sámi soagknegka* [Small North Saami-Kildin and Kildin Saami-North Saami dictionary. Ohcejohka: Girjegiisá.

Sharshina, Nina S., Elizabet Sheller [Elisabeth Scheller] and Aleksandra A. Antonova 2008. Sām'kilsyjjt (ōxxpnuvvemtyj): 1. grammatika: 2. leksika. Unpublished script. Lujávv'r.

Scheller, Elisabeth 2004. *Kolasamiska – språkbyte eller språkbevarande? En sociolingvistisk studie av samernas språksituation i Ryssland*. M.A. thesis. Umeå: Umeå universitet.

— 2011. Samisk språkrevitalisering i Ryssland – möjligheter och utmaningar. *NOA:* 27 (*1*): 86–119.

— [in print] 2012. Kola Sami Language Revitalisation – Possibilities and Challenges. In *L'Image du Sápmi 2*. K. Andersson (ed.). Humanistica Oerebroensia. Artes et linguae. Örebro universitet.

Sergeeva, Elena 1995. The Situation of the Sami People in Kola. In *Indigenous and tribal peoples' rights – 1993 and after*. E. Gayim and K. Myntti (eds.), 17–188. Juridica Lapponica 11. Rovaniemi: Institute for Environmental and Minority Law.

— 2000. The Research History of Kola and Skolt Sami Folklore and Religion. In *Sami Folkloristics*. J. Pentikäinen et al. (eds.), 155–188. NNF Publications 6. Åbo: Åbo Akademi.

— 2002. The Eastern Sámi Languages and Language Preservation. In *Samiska i ett nytt årtusende*. E. Mørck and T. Magga (eds.), 103–116. Vol. 717. ANP. Köpenhamn: Nordiska Ministerrådet.

Siegl, Florian and Michael Rießler (in print). Uneven Steps to Literacy – the History of Dolgan, Forest Enets and Kola Saami Literary Languages. In *Equally Diverse: Comparing Linguistic and Cultural Minorities in the Russian Federation and the European Union*. Reihe: Uralica Helsingiensia. H. Martin, J. Saarikivi, R. Toivanen and M. Rießler (eds.). Helsinki: Helsingin yliopisto.

Trosterud, Trond 1996. *Funny Characters on the Net: How Information Technology may (or may not) support Minority Languages*. BAR-IT conference, Apatity, 16.-

20.9. 1996. http://www.hum.uit.no/a/trond/barit.html.

— 2001. Review of Kåven, Brita E. (red.) 2000: Stor norsk-samisk ordbok / Dáru-sámi sátnegirji. Davvi girji, Kárášjohka. In *LexikoNordika* 8: 283–306.

Yur'ev, Aleksandr 2003–2007. *Elektronnyi saamskii slovar' = Electronic Saami dictionary.* RuSaDic. Ver. 2.0. CD-ROM.