Lexicography and indigenous languages

Ole Henrik Magga, Sámi allaskuvla/Saami University College, Guovdageaidnu/Kautokeino

Keywords: indigenous languages, Saami, lexicography, language planning, knowledge in language.

Abstract

Indigenous languages are disappearing rapidly. The UN system and especially UNESCO supports the work with indigenous languages for several reasons. Language death is not a loss only for indigenous peoples themselves, but for all humanity because the loss of languages affects diversity, and this diversity is essential for our survival. Lexicographical work is essential for indigenous languages, particularly pursuant to their use in education and society at large. All Saami languages are endangered languages. The lexicon of almost all the ten Saami languages has been investigated quite thoroughly after the first dictionary was printed in 1738. Konrad Nielsen’s dictionary published in 1932-62 is the masterpiece within Saami lexicography. Dictionaries have also been compiled for educational purposes and within specialized fields, like anatomy and mathematics. Embedded in the vocabulary of the Saami languages is knowledge from the past over 6000 years back in time and the loanwords give much information about contacts with Balts, Germanic peoples, Scandinavians and Russians. Saami languages have a rich descriptive terminology and terminology on nature and animals, especially on reindeer. Saami is probably the richest language on snow terminology in the world. For the past 40 years substantial resources have been allocated, especially in Norway, to the education of teachers and printing of books, including dictionaries, in Saami. This would not have been possible without the lexicographical research that had been carried out at the universities throughout a number of years. Dictionaries and grammar books are the cornerstones of language teaching and of language use, as they provide documentation and organization of information about language structure and the units of language.

1. Introduction

There were 6809 languages in the world in the year 2000 (UNESCO 2003: 24). There is reason to believe that a majority of these were indigenous languages. But the figure is certainly not 6809 anymore. UNESCO has estimated that if nothing is done, half of these languages will disappear by the end of this century. With this, 364 may have died out since 2000, and 6445 may have survived. This is a guess based on
statistics, but probably not far from the truth, sadly enough. Since 1982 indigenous peoples have worked together with each other, with experts, with governments and the UN system in order to formulate principles aiming at securing indigenous cultures and languages a chance to survive and develop. And finally, in 2007, a *United Nations Declaration on the Rights of Indigenous Peoples* was adopted the General Assembly. Article 13 (first paragraph) has this wording:

Indigenous peoples have the right to revitalize, use, develop and transmit to future generations their histories, languages, oral traditions, philosophies, writing systems and literatures, and to designate and retain their own names for communities, places and persons.

UNESCO sees links between the earth’s linguistic, cultural and biological diversity. About the loss of languages, UNESCO has a very clear view: “With the disappearance of unwritten and undocumented languages, humanity would lose not only a cultural wealth but also important ancestral knowledge embedded, in particular, in indigenous languages.” (http://www.unesco.org/new/en/culture/themes/endangered-languages/, 2012-06-12). Even if this insight comes late, it is still very encouraging.

Like other parts of traditional knowledge, indigenous languages have been transmitted from generation to generation orally without documentation in the form of writing. This has also been my own personal experience. I was 22 years old when I had the privilege of being taught by a teacher, for the first time in my life, how to write the words I had been using since childhood. It happened in Oslo and I also learned how to look up words in a dictionary. Since then I have learned much about why this happened to me and why and how it has happened to many other indigenous individuals. Many of them still have no chance to learn how to write their own language. Whole communities have already lost their language and with it much of their knowledge. It is not only a loss from them, but a loss of parts of the total ‘pool of ideas’ or the universe of thought produced by human beings. The UNESCO publication I referred to initially has described the value of languages in this way: “Linguists and anthropologists have suggested that the diversity of ideas carried by different languages and sustained by different cultures is as necessary as the diversity of species and ecosystems for the survival of humanity and of life on the planet” (UNESCO 2003: 30). The contents
and design of dictionaries has immense value for indigenous cultures, for language learning and teaching, for history, ethnology, land rights – for all kinds of knowledge that can be based on, derived from, or related to data about words, their meaning, form and history. Many indigenous languages have not been explored at all yet. We need and we welcome all kinds of linguistic work, and in particular lexicographic work. Based on the experiences we have had with this kind of work in our languages, I will present to you why we see lexicographical work as very important.

2. The Saami languages

Traditionally there are ten Saami languages, the number of speakers in parenthesis:

1. South Saami (300-500)
2. Ume Saami (few old speakers)
3. Pite Saami (few old speakers)
4. Lule Saami (2000-3000)
5. North Saami (17 000)
6. Anár Saami (300-500)
7. Skolt Saami (300-500)
8. Akkala Saami (extinct)
9. Kildin Saami (650)
10. Ter Saami (few old speakers)

Endangered (UNESCO Atlas):
Definitely: 5
Severely: 1,4,6,7,9
Critically: 2,3,10
Extinct: 8

The Saami language group is a continuum in the sense that the phonological forms of word stems may easily be converted from one language to another. They belong to the Finno-Ugric branch of the Uralic language family. Finnish as their closest sister language together with the other languages of the Baltic-Finnic group: Estonian, Karelian, Veps, Ingrian, Votic, and Livonian. Other languages in the Finno-Ugric branch of this family are the minority languages Mordvin, Mari, Udmurt, Komi, Mansi and Khanty in Russia and Hungarian. Nenets, Enets, Ngasar and
Selkup in Russia make the Samoyed branch of the family. The Uralic proto-language is believed to have existed 6000 years ago.

3. Overview of Saami lexicography

The Saami lexicographical work started in 1557, literally by accident. After a shipwreck near the cost of the Kola Peninsula the English captain Stephen Borrough had to stay some time in a Saami community. His list of 95 words was published 30 years later in *The Principall Navigations Voiages and Discoveries of the English Nation*. It then took almost 200 years before the first real dictionary was made, in Sweden by the priest Pehr Fjellström, in 1738. Since that time Saami lexicography has made much progress and it is impossible to mention all the published works and their contents. However, they fall quite naturally in four categories:

- dictionaries intended for priests and missionaries
- dictionaries for scientific purposes intended as documentation of language and culture
- dictionaries intended for education institutions and the general public
- specialized dictionaries (LSP dictionaries) within one or more subject matters

3.1. Dictionaries intended for priests and missionaries

Fjellström’s dictionary was explicitly intended for teachers and priests. It was an instrument in the very active Lapland policies pursued by the Swedish kings starting already with Carl IX around 1600. His first project was to kidnap ten Saami boys and try to bring them to Uppsala so they could become priests. It did not work. They all escaped before they reached Uppsala. A better outcome was achieved with the printing of books and building of schools and churches. The peak achievement within lexicography was the grand dictionary *Lexicon Lapponicum* by Lindahl and Öhrling in 1780. It was based on Ume and Pite Saami, even if the spelling standard was called “the South Saami book language”, in which the whole Bible was printed in 1811.

From the beginning of the 18th century the Danish-Norwegian kings had taken up the challenge from the Swedes and Russians in the
race for Lapland and its resources. Missionary work was the most important instrument also for them. A *Seminarium Lapponicum* was established in Trondheim in 1752 with the aim of educating teachers. Professor Knud Leem was the headmaster. Already in 1756 he published his first dictionary. In 1769 the first part of his voluminous dictionary *Lexicon Lapponicum bipartitum* (Saami-Danish-Latin) was printed and part nr 2 twelve years later in 1781. The reason for this delay was the Venus passage in 1769. Along with the astronomer Maximillian Hell came a Hungarian linguist named Sajnovics, who after this visit claimed to have demonstrated that Hungarian and Saami was “the same language” and should be spelled in the same way, which caused a long dispute with Leem on orthography. Leem died in 1775 and did not see the printing of the second part of his work. Leem was a remarkable linguist for his time. His works both with Saami and Norwegian dialects were very accurate and he used only material which he had collected himself, while Lindahl and Öhrling used much material of obscure origin. Still, the academic world has continued referring to Lindahl and Öhrling’s dictionary much more than Leem’s.

In the course of the 19th century, the former officer in the Danish army, Niels Vibe Stockfleth, who had fought in the Napoleon wars, became the leading researcher, teacher and missionary in North-Norway. He published his *Norsk-lappisk Ordbog*, 1851 (Norwegian-Saami). Even if Stockfleth’s work was groundbreaking, especially with regard to orthography, which he developed in cooperation with the famous Danish linguist Rasmus Rask, the work of his student and successor Jens Andreas Friis was much more accurate, resulting in 1887 in a Saami-Norwegian-Latin dictionary with 868 pages titled *Ordbog over det lappiske sprog*. These two created a new spelling of North Saami with the special “Saami letters” š, č, ž, ѣ, t and þ, and many people learned to read Stockfleth’s and Friis’ publications and especially the Bible translation from 1895.

### 3.2. Dictionaries for scientific purposes and documentation

Finno-Ugric and Uralic studies were established in the course of 19th century and in that spirit much work was done even with the Saami language. An important impetus to this was the Finnish research in history, language and culture which started at that time as part of the nation building process that eventually resulted in the Finnish state. Also Hungarian, Swedish and Norwegian researchers made very important
contributions within Saami lexicography. Already in the 1820s the first record was made of South Saami words. From 1880 the Hungarian Ignácz Halász, the Finn Arvid Genets, the Norwegian Just Qvigstad and the Swede Karl Bernhart Wiklund were the first to document vocabularies from several Saami languages. They were as a rule all written with phonetic transcription and translated into German, which was the Fachsprache in this field for a long time. In the course of the next hundred years much lexicographical work was carried out within this tradition. Halász published his material, including vocabularies in a series of works titled Svéd-lapp nyelv I-VI 1885-96. Wiklund’s Lule-lappisches Wörterbuch (Saami-German) was published in Finland in 1890. Arvid Genetz: Kuollan lapin murteiden sanakirja ynnä kielennäytteitä/ Wörterbuch der Kola-lappischen Dialekte nebst Sprachproben (Saami-Finnish-German) came in 1891. The Finnish linguist Eliel Lagercrantz made a remarkable contribution by collecting texts, writing grammars and publishing two very important works at that time, Wörterbuch des Südlappischen (1923) and a unique compilation of material from many Saami languages, Lappischer Wortschatz I-II (1939). T. I. Itkonen documented the Kola Saami and Skolt Saami vocabulary in an extensive publication in phonetic transcription and Finnish and German translations titled Koltan- ja kuolanlapin sanakirja 1-2 (1958). Harald Grundström published a Lule Saami dictionary in four volumes, Lulelappisches Wörterbuch (1946-54) with German and Swedish translations, based on his own material and material collected by Wiklund, Collinder and others. The lexicon of Ume Saami was studied by Wolfgang Schlachter resulting in Wörterbuch des Waldlappendialeets von Malå and Texte zur Ethnographie (1958). Gustav Hasselbrink carried out an enormous compilation of all previous published material in South Saami resulting in Südlappisches Wörterbuch = Oårj’elsaamien baaguog’ärjaa (1981-1985). He rewrote examples into his own phonological transcription and made translations in German. An Anár Saami dictionary has been compiled by Erkki Itkonen, an important and the most productive researcher within Saami historical linguistics, titled Inarilappishes Wörterbuch I-IV (1986-91).

All these works mentioned have their value for anyone who works with these languages scientifically. For a broader public however, there are in general two aspects that limit their value. Phonetic transcription is not well known among ordinary people and even broader transcriptions can be difficult to read. The access for ordinary readers is even more reduced by the translations, especially when they are only into German.
The masterpiece within this tradition is the five volume *Lapp dictionary* (1932-62) by the Norwegian scholar Konrad Nielsen, who was a professor at the University of Oslo. It was based on Nielsen’s thorough scrutiny of the vocabulary in three main North Saami communities Deatnu-Tana, Kárášjohka-Karasjok and Guovdageaindu-Kautokeino. His phonological analysis represented a breakthrough and is still valid today, and his translations into Norwegian and English are very accurate. His dictionary and orthography is still used by scientist in this field. It is still “the dictionary” for all users of North Saami, even if the spelling has been changed twice since its first parts were printed. It gives a very accurate description of both form and meaning of North Saami lexicon and it has played a key role for teaching, research, language planning for 70-80 years. It was reprinted in 1979 by common effort from many institutions in Scandinavia.

Numerous word studies have been carried out. The grand old man in lappology J. Qvigstad published his *Nordische Lehnwörter im Lappischen* in 1893, and N. E. Hansegård, an important scholar in research within bilingualism, published his *Recent Finnish Loanwords In Jukkasjärvi Lappisch* in 1967. Juhani Lehtiranta then with his *Yhteissaamelainen sanasto* (Common Saami vocabulary) 1989 demonstrated that at least 1479 word stems are common to all Saami languages.

3.3. *Dictionaries intended for education institutions and the general public*

Soon after the borders were settled in Saamiland in 1751 and 1826, a combination of Norwegian national romanticism, a fear of Russian and Finnish hegemony and vulgar-Darwinian ideas led to the launch of a massive assimilation campaign towards the Saamis and immigrated Finns, also called Kvens. The Saami and Kven languages were forbidden from being used in schools from 1890s to around 1970, when Saami was reintroduced into education. In Norway and Finland, recent political reforms in the 1980s and 1990s led to recognition of the Saami as a distinct people. In 1988 the Norwegian Parliament passed a constitutional amendment adding a new section 110A to the constitution with this wording: "It is incumbent on the governmental authorities to take the necessary steps to enable the Saami population to safeguard and develop their language, their culture and their social life." Similar amendments
have been made to the Finnish constitution, while the Swedish constitution only recognizes ‘minorities’ without special mention of the Saami people. In 1992, a language legislation was passed in Norway and Finland and in the year 2000 the same took place Sweden. For the past 40 years, substantial resources have been allocated, especially in Norway, to the education of teachers and printing of books, including dictionaries. This would not have been possible without the research that had been carried out at the universities, even during the “cold winter”. Many dictionaries have been compiled for educational purposes, as well as for the general public, with translations into and from Finnish, Norwegian, Swedish, Russian and German. Here are some of them (the numbers refer to the list of Saami languages):

4. Specialized dictionaries (LSP dictionaries) within one or more subject matters

As Saami has gained some official status at least locally, the demand for terminology and dictionaries in a variety of fields has increased. In several fields, word collections and vocabularies have been printed, some of them as a part of language planning work:


Like Saami, most indigenous languages have been kept outside of the modernization of society that took place during the 20th century. Introduction of indigenous languages into the education system and society is therefore for many reasons far more complicated than most people seem to understand. The mentioned works from scientists from various countries have demonstrated the value of academic work at its very best. Almost all the dictionaries in category 3 and 4 are based on earlier research and dictionaries. Saami is really in a very fortunate position compared to many other indigenous languages. Several of the Saami languages can be used for teaching and publication purposes on highest level without much difficulty. Sámi allaskuvla/Saami University College where I work, carries out both education and research in Saami. Lexicographical work has been and is a meaningful activity that we can not do without.

4. Knowledge embedded in indigenous lexicography

Indigenous peoples are often said to have no history. And it is true that majority historians for a long time did not pay much attention to our history. (In present times also this is changing). We have scarce written
documentation from the past. UNESCO has emphasized the role of indigenous languages as reservoirs of ancestral knowledge. And UNESCO is right about this. In language, there is much information, even from far back in time.

In Saami society, family relations and hence kinship terminology are important. There is a distinction between words for uncles and aunts based on their age in relation to ego’s father and mother. Mother’s elder sister is gāaskie and younger sister is muahra. Father’s younger brother is tjietsie and elder brother is jiekie (the forms are South Saami). Both gāaskie, tjietsie and jiekie derive from Proto-Uralic 6500 years ago. ‘Daughter-in-law’ is mænnja and ‘younger sister-in-law’ (younger than ego’s wife) is nāāte in South Saami. All date back to this period. Family relations of today had already the “same” terms 6500 years ago. English terms like daughter-in-law seem to be very young compared to this. From this time we can follow the development in several fields.

Copper was used 6500 years ago and called something that today is veaiki in North Saami. The Sami learned about selling from a Proto-Indo-European people (doekedh). 5500 years ago our foremothers and -fathers had reindeer (bovtse) and dogs (bïenje), they had nets (soaje ‘net needle’, voektenje ‘bar for drying nets’). 5000 years ago they had boats (vïnhtse) with a steering oar (North Saami mealli), and they also they needed to pay (maeksedh) (Sammalahti 1998).

Recent research in Saami lexicography has even uncovered possible traces from a substratum from pre-Saami time. One example is the element -ir which may have had the meaning ‘mountain’ (Aikio 2009). When it comes to salmon and hay (loese, suejnie), this was probably learned from the Finns, who in turn had learned about these from a Proto-Baltic people. The words saertie ‘heart’ and liejpie ‘alder’ indicate direct contacts with the Balts.

The contact with Germanic peoples was probably established before contact with Scandinavians. Words like ebrie ‘rain’ buejtie ‘fat’, boelte ‘slope’ and laajkoe ‘rent’ suggest this. Farm animals like sheep (saavtse), cow (govse) and goat (gaahtse) and to milk them (mielhkie ‘milk’) the Saami learned from Scandinavians, and the same was the case with metals like gold (gullie), silver (sïlpe), tin (ditnie) and iron (ruevtie).

Also Russian has contributed to the vocabulary, directly as is the case with ohpit ‘again’ and through Finnish or Karelian as demonstrated by North Saami bähppa ‘priest’ (Fi. pappi) and raejje ‘border’ (Fi. raja).

The Proto-Nordic loans in Saami (and Finnish), and later loans from Nordic, have been widely discussed, and have proven to be very
useful tools in the work of dating and confirming historical events in the past. The word for scissors *skaarja* was borrowed before the change *a>e* took place in Nordic (i-umlaut). In the place-name *Máhkarávju* (name of the island where North Cape is situated), *ávju* reflects the Proto-Nordic n form of Norwegian φy ‘island’, thus attesting early Norwegian presence in the north. The name of the river Namsen in Trøndelag, South Saami *Laakese* seems to reflect a proto-Nordic *laguz*, which is a strong evidence of early Saami use of the land in the south and therefore of much value in the on-going debate on issues of Saami land rights in Norway and Sweden.

There are Saami loanwords in other languages, most in Finnish with 373 lexical items, most of them words relating to reindeer herding (24 %) and topographical terms (17 %). (Aikio 2009). A few loans are found also in Swedish and Norwegian, like *rev* ‘fox’ and *jokk* ‘river. Our international contribution is *tundra*, a word that obviously originates from Kola Saami and has passed on to other languages through Russian. Also *morse* comes from Saami.

Several studies on terminology have given valuable information about the past (Korhonen 1981, Sammallahti 1999, Nesheim 1967). The points of contact with the outside world throughout history are nicely attested by the distribution of loanwords on a map, like this for *wadmal* (woven wool fabric):

(from Nesheim 1967)
The physical environment has an influence on cultures and languages. The Saami culture has had a long and intimate relationship with the arctic environment. The languages have terminologies for nature, human beings, animals (especially reindeer), and snow and ice. A male reindeer have names based on age from 1 to 7 years in this sequence beginning with calf: miesse, čearpmat, varit, vuorsu, gottos, goasohas, máhkan, nammalahppi (North Saami). Body size, shape, condition, color, antlers, nature of the coat and individual personality can be described accurately with hundreds of terms (Magga 2006). Nielsen’s dictionary contains detailed accounts of terminologies on all these fields, and they were even classified according to meaning by Asbjörn Nesheim in the latest volume.

Already at introductory courses in linguistics, the students are told that indigenous peoples in rainforests have both terms and knowledge about plants and animals that are important for discovery and development of new medicines and drugs. All this is true. And we, the Saami, are in competition with our good neighbors, the Inuits, when it comes to snow and ice terminology. It began with the famous anthropologist Franz Boaz’ documentation of a few words from Inuit, which were picked up by the even more famous Benjamin Whorf, who referred to the number of snow words as “two hundred”. Ivan Krupnik at the Smithsonian Institution and Ludger Willer-Mülle have settled this misunderstanding in a wise way: The Inuits are still the ice people, but when it comes to snow they write: “So, if some linguists or journalists are interested in counting someone’s “words for snow”, we have a message for them. Please switch to another language! The Norwegian Saami, who tend their reindeer herds over the northernmost realm of Europe, reportedly have 100 words for snow (Magga 2006)” (Krupnik et. al. 2010: 403). Inger Marie Gaup Eira defended her PhD-dissertation The silent language of snow in June this year, reporting over 300 terms from Guovdageaidnu that have to do with snow. Here is her classification of some of the terms:
And North Saami is not unique in this respect among the Saami languages. South Saami is extremely rich on terminology on layers of snow and ice. These are the terms on crust types on the top of the snow pack, beginning with the thinnest layers and ending up with hard ice:

- **skaevie** first trace of crust on snow,
- **tråapkedh** thin crust beginning to form (in the spring in the evening) that does not bear skis,
- **skrâehrie** thin crust,
- **skaerrie**, crust (on snow),
- **skåerie**, fine crust on snow (that creaks when you walk on it),
- **snjaefie** very thin crust so that it is barely possible to go skiing and the snow is not sticky (after thaw),
- **snjâehrie** thin crust (in the evening in the spring),
- **spieltie** ice crust (from frozen rain),
- **tsaapke**, crust that does not bear,
- **traapke** weak crust so that the skis sink through occasionally,
- **tsievie** snow so hard that it bears small reindeer,
- **radte** the kind of going that the crust (after thaw, especially in the autumn) destroys pasture,
- **slaptjedh** crust that is beginning to thaw,
- **tjarve** crust (that bears),
- **gielper-tjarve** hard crust,
- **guelpere tjarve** very hard crust,
- **njahpedihkie** crust hard as ice.

Researchers working with international snow and ice terminology have already for some time been working together with Saami linguists in order to learn from natural terminologies on ice and snow (Jernsletten 1997). The Saami languages are also rich in terminology on position and movements in a landscape, and descriptions of appearance, states and movements of animals and human beings.
All languages need new terminology and especially languages that have been used within limited domains. Saami has a rich morphology and many possibilities for making derivations. A prerequisite for making new lexemes by derivation is to know existing morphological and semantic patterns of the language. If one needs a term for the result of an act expressed by a verb, or the instrument for that act, in South Saami, one can look at derivation patterns like this, which are common knowledge to lexicographers: mojhtese ‘memory’ < mujhtedh ‘remember’, vadtese ‘gift’ < vedtedh ‘give’, damtese ‘acquaintance’ < damtedh ‘know’, gaptjese ‘covering’ < gaptjedh ‘cover’, doengese ‘patch’ < doengedh ‘patch, mend’. A new term ending on -ese can easily be derived from many verbs. Dictionaries are in this way extremely useful tools for language planning, and every day they are used for this purpose.

5. Challenges

The most important task for indigenous lexicographical work is to get recordings of speech before it is too late. Lexicographical work in itself can be complicated, and when there are only a few speakers left of a language, it becomes very tough. From the work with a South Saami dictionary as an assistant to my wife, I know how time consuming it can be to check all details from all kinds of sources like archives, books, tapes and from those few older people who can serve as a reliable source (not everyone can) – and one still ends up with doubts. Some of the communities from where the information was collected, do not exist anymore. So we often end up with piecwise description of words from different geographical areas from different times. We were often completely dependent on the original sources, and frequently we were not sure whether two words really were different in meaning or not. The word bovresjidh had been translated with ‘make a hole with a knife’ in the old sources and the word bàerihtidh ‘drill a hole with a knife’, but nobody could tell if they in fact were identical. We wanted to be loyal to the sources and ended up with two different translations. In addition you also have to modernize the translations, and this can also be a source of inaccuracies. Words may even have surprisingly different meanings like the word vaejsjie that has been documented as having the meaning ‘beast of prey’ in some sources, and ‘elk, moose’ in others.
6. Conclusions

Lexicographic work is extremely important for indigenous languages. We are grateful for what has been done in the past and we are full of expectations for what more lexicographers can do, both when it comes to Saami and with other indigenous languages. Work with indigenous languages is rewarding because there are so many phenomena that have still not been properly described and understood in languages. My teacher professor Knut Bergsland, who used his knowledge to help local Saami communities substantiate their claims of land rights, often said that Chomsky should have studied indigenous languages more before launching his TG-theory. So they even are of interest to general linguistics. Dictionaries and grammar books are the cornerstones of language teaching and of language use, as they provide documentation and organization of information about language structure and the units of language. The best things we can do in order to preserve and develop languages, is to use the languages, and to enable others to use the languages.

References
