

Nordic language ambassadors - Teaching speakers of small Nordic languages about language technology

Marie Mattson

The Language Council
The Institute of Language and Folklore
marie.mattson@isof.se

Abstract

Here, I describe a part of the Nordic initiative for raising awareness and competence of language technology for the small languages of the Nordic countries. For many small languages, there is not enough language technology to support digital communication and infrastructure. In the cases where sufficient language technology has been developed, not enough people know about it or how to use it. Therefore, the Norwegian Ministry of Education and Research has instructed the Swedish Institute of Language and Folklore (ISOF) to teach end users how to use existing language technology. We did this by finding language ambassadors for some small languages who in turn teach their language community about language technology.

1 Background

Today, hundreds of languages are spoken in the Nordic countries, including the national languages (e.g. Swedish, Danish and Norwegian), the national minority languages (e.g. Sami, Kven Finnish and Romani). This project is an effort to raise awareness of the importance of language technology (LT) in the language communities to help the languages survive in the increasingly digital world.

1.1 Language technology for the small languages in the Nordic countries

Today, it is necessary to have access to technology in order to pay the bills or book a doctor's visit. This works well for the majority languages, but not for many of the small languages. Speakers of the small Nordic languages often have insufficient accessibility to technology in their language and have to get by with one of the majority languages instead. This leads to digital exclusion, which becomes more serious the more digitalized the world becomes (Domeij et al. 2019).

The different Nordic languages all have varying amounts of accessible language technology. The Nordic language technology group, ASTIN, explained the situation for the small, Nordic languages in the European Language Equality project (Nørstebø Moshagen et al. 2022). Some small languages, like the Romani languages, Karelian and Pite Sámi lack basic LT resources. However, for others such as Faroese, North Sámi and Lule Sámi, there are several LT tools available, even advanced tools like speech synthesis or machine translation in a few cases. Operating systems often recognize these languages, but the LT tools are often not widely used by or available for speakers. There are also languages that fit somewhere in between the aforementioned categories, like South Sámi, Kven and Meänkieli, which have access to some basic LT tools on some platforms.

Not having access to sufficient LT is a threat to small languages in today's digital society. Many speakers of small languages are bi- or multilingual which makes them less likely to use the language in writing which is a big threat for their existence. This digital exclusion can even lead to language extinction if it continues (Borin, Brandt, Lachica Parkvall 2012).

1.2 The Nordic initiative – inclusion of small languages in the Nordic countries

Because of the aforementioned challenges the small Nordic languages are facing, ASTIN and The Arctic University of Norway (UiT) created a Nordic LT initiative. The Nordic initiative is separated into three parts.

Part 1 of the initiative (Swedish): *Nordisk insats för att öka kunskapen om språkteknologins betydelse både för ökad digital inkludering och för att stärka och skydda de små språken i Norden.*

English (unofficial translation): *Nordic measure to raise awareness of the importance of language*

technology in order to reach a higher level of digital inclusion and to protect and strengthen the small languages in the Nordic countries.

The other parts of the Nordic initiative aims to:

- include the small languages in the EU's work with linguistic equality by 2030¹.
- influence the tech-giants to include small languages in the Nordic countries.

The Nordic Language Technology group, ASTIN, gave several recommendations for improving the situation for the small and low-resource languages. One of the recommendations is to “raise awareness in the language communities of the importance of language data and the availability of such data for language technology” (Nørstebø Moshagen et al. 2022:42), which this project is aiming to do.

2 The languages and language ambassadors

We were not able to include every small language in this project. Instead, we focused on languages that had some LT tools that we could teach to end users. The following languages were included in the project:

- Languages that have access to basic tools on all platforms:
 - North Sámi
 - Lule Sámi
 - Faroese
- Languages that have access to the basic tools on most platforms:
 - South Sámi
 - Meänkieli
 - Kven Finnish²

All of the languages are national minority languages in Sweden, Norway or Finland with the exception of Faroese, which is an official language on the Faroe Islands.

All the languages have LT tools developed by the Arctic University of Norway in Tromsø (UiT).

¹One example of this is the European Language Equality project: <https://european-language-equality.eu/>

²Categorization is based on the one made by Astin in the European Language Equality report (Nørstebø Moshagen et al. 2022:41)

Although some languages only have a few LT tools that are available to the public, UiT (among others) are actively working on developing more tools for these and other languages.

2.1 Language ambassadors

In order to reach the end users, which was our final goal, we recruited one language ambassador for every language in the project. I was looking for people who work with one of the languages and therefore already have a network of people who might be interested end users. They also likely have access to varying output channels that already have an audience (for example a website, social media and in some cases language center visitors and students).

To find language ambassadors, I contacted different people and organizations who work with the language in question. I contacted universities, language councils, the different Sámi councils and language institutes and centres in the Faroe Islands, Norway, Sweden and Finland.

I managed to find one ambassador for every language. They were from the Faroe Islands, Sweden and Norway. They were mostly language experts at language centers or language institutes, but there was also one teacher and one master's student in the Faroese language. Aside from the official language ambassadors, there were four additional participants, two who worked with Kven Finnish and two who worked with North Sámi.

3 Educating the language ambassadors and end users

We educated the language ambassadors in two steps, first with Giellatekno from UiT Tromsø, and then a workshop in Copenhagen held by the Language Secretariat of Greenlandic. After this, the ambassadors will hold their own workshops targeted at end users.

3.1 The language technology course and workshop

The first step of educating the ambassadors was to visit The Arctic University of Norway in Tromsø. Here, Trond Trosterud held a weeklong course about which LT tools³ are available, how to use them, what to do when they do not work as intended and how to teach this to end users.

³The LT tools the ambassadors learned about are available here: <https://giellatekno.uit.no/> and here: <https://divvun.no/>.

Some of the LT tools they learned about were spell and grammar control, mobile and computer keyboards, machine translation and text-to-speech.

In November of 2022, Per Langgård, Beatrine Heilmann and Judithe Denbæk of Oqaasileriffik, the Language Secretariat of Greenlandic, will hold a workshop about Greenlandic LT in Copenhagen. Greenlandic LT has come very far in its development despite it being a small language. Among other things, the ambassadors will learn about why the Greenlandic LT have been so successful. This is meant to inspire them and give them an idea of what their language community can do to develop and spread LT successfully.

3.2 The ambassadors' workshops to reach end users

The ambassadors have not held their workshops aimed at end users yet. Therefore, in this section I will present a few examples of some of the plans and ideas that the group has discussed.

Because the ambassadors know their language community the best, we entrust them with deciding what measures will be most impactful for their community. ISOF and Giellatekno and Divvun from UiT give any support the ambassadors may need. Their trips and venues are paid for in order to let the ambassadors focus on the content of their own workshops.

Throughout the project, we have suggested the ambassadors to host *workshops*, but it does not necessarily have to be workshops. Activities like speaking at a conference or making a guide or a video about the LT tools are also options that some of the ambassadors seem to be interested in doing.

4 Summary and hopes for the future

To summarize, we have found ambassadors for six of the small languages of the Nordic countries. They have been on a week long course at UiT where they learned about what LT tools are available at the moment and how to use them. The next step in this project is for the ambassadors to go to Copenhagen to learn about Greenlandic LT and then they are going to plan their own workshops.

So far this has been a successful project and is a great example of cooperation between the Nordic countries. We believe similar projects could lead to positive outcomes for other small languages. The project would also benefit from including Finland in some way. We have seen that there has been an

interest for the project in the language communities which is proven by the fact that we got four more participants than we first anticipated. There also seems to be an interest in the language communities because both Kven media (Vaara 2022) and Meänkieli media (Niiva, Kangas 2022) have mentioned this project in an article and on radio.

The project has been a great way to spread knowledge about LT to people who speak and work with the small languages of the Nordic countries. Other than that, another impactful aspect of this project has been establishing a network between ISOF, UiT, the Language Secretariat of Greenland, the Sámi Councils and the other language institutes. It is crucial that we keep in touch with this network in the future if we want the interest and quality of LT for small languages to continue growing.

References

- Borin, Lars, Brandt, Martha D., Edlund, Jens, Lindh, Jonas, Parkvall, Mikael (2012). The Swedish Language in the Digital Age/Svenska språket i den digitala tidsålder. *META-NET white paper series*.
- Domeij, Rickard, Karlsson, Ola, Trosterud, Trond, Nørstebø Moshagen, Sjur / [ed] Sullivan, Kirk P. H. and Coppélie Cocq, Leiden (2019), Enhancing information accessibility and digital literacy for minorities using language technology: The example of Sami and other national minority languages in Sweden (Indigenous Writing and Education), *Perspectives on Indigenous writing and literacies*, doi = "10.1163/9789004298507"
- Niva, Henrik, Heikka, Maria, Kangas, Elina (2022), Meänraatio Gränslöst med Henrik Niva och Maria Heikka 9 august, <https://sverigesradio.se/avsnitt/granslost-henrik-niva-11>
- Nørstebø Moshagen, Sjur, Domeij, Rickard, Eide, Kristine, Juel Henriksen, Peter, Langgård, Per (2022), Report on the Nordic Minority Languages, doi = "10.1163/9789004298507"
- Vaara, Tomi (2022), Unike språkverktøy brukes for å styrke små nordiske språk, *NRK Kvensk/NRK Kvääni*, <https://www.nrk.no/kvensk/unike-sprakverktoy-brukes-for-a-styrke-sma-nordiske-sprak-1.16094738>