

How open is open - visibility and accessibility from a Greenlandic perspective

Per Langgård

Oqaasileriffik/ The Language Secretariat, Nuuk, Greenland

per@oqaaserpassualeriffik.org

Abstract

Language technology for Inuit languages is vital for language survival. On the surface it should be easily provided since (i) linguistic rights for most of the Inuit dialects are well secured, (ii) Inuit languages maintain a very high status among its speakers, (iii) the need for technological solutions is recognized at the political level and (iv) funding for projects on Inuit culture and language is comparatively easy to obtain. Still, only one working project is found.

A number of reasons for this state of affairs will be identified and a case made to show that an extremely easy access to all kinds of free resources is the only option for Inuit languages to enter into the much needed world of language technology

We are doing well

The Greenlandic language technology project is not very old. Neither is it very big in terms of staff or other resources, and academic achievements are very meagre this far.

Still, the project has attracted vast amounts of attention not only from lay Greenlanders but we have also noted quite a lot of interest among professionals in the field.

We are of course very pleased to see that our efforts pay off and very proud whenever we hear mention of our project in academic circles which we do comparatively often, basically for two reasons:

(i) Language technology programmes for Inuit languages apart from a few attempts conducted by southern scholars are non-existing.

(ii) Greenlandic is notorious for morphological complexity. Until we launched the first finite state automaton in 2006 the standard attitude was a total rejection of language technology for Inuit languages even among the most prominent scholars in the field. A polysynthetic language cannot be computerized. I could add that I even today come across high ranking linguists in the field of Eskimology who maintain that Mother Earth is flat.

We are of course proud and happy to collect the laurels to Greenland but it would be very hypocritical to leave it there, for without computational linguists and computational

scientist with a serious wish to share their own achievements with the rest of the world we simply would not be where we are. In our case the guardian angels are situated in Tromsø/ Kautokeino and in Odense, but it could no doubt have been Gothenburg, Oslo, or someplace else had things developed just a bit differently back then in 2005 when the whole project started.

So congratulation Tromsø and Odense, and congratulation to all the rest of you who believe in open resources. We did it together!

Analysis

It is a fact that the need for language technology to support minority languages - especially threatened ones - is generally recognized in the political bodies with direct influence and power like ICC and RAIPON. The Tromsø conference on indigenous languages in the Arctic in October 2008 is an obvious example.

Still, the attempts seldom make it past declarations of intent into concrete projects or good actual projects soon dry out and die. Greenlandic language technology is no doubt one of the few real sunshine stories of its kind, not only among Inuit languages but also among minority languages as such. I would therefore like to take the opportunity to address two of the questions this statement gives rise to at this point:

(i) How come it is so difficult?

(ii) What will it take to pave the way for many more projects like the Greenlandic one?

The analysis to follow is primarily based on my experiences with Greenlandic, Iñupiaq, and Inuktitut but rather many encounters with representatives for American First Nations and RAIPON (Russian Association of Indigenous Peoples of the North) have given me the impression that the observations hereunder are much more widespread than I originally believed them to be.

As a very first answer to the question above most of us will no doubt think in terms of unrecognised linguistic rights, lack of financial and linguistic resources, or the like.

This is no doubt part of the truth and thirty

years ago I would myself gladly have accepted it as the whole truth. But not any more. The situation is simply far more complicated than we believe it to be for an immediate glance. Linguistic rights have been only partially threatened in Greenland and only for a very limited number of years. Greenlandic as an official/ recognized language has a long history. From the very onset of mission in 1721 Greenlandic speaking Greenlanders have had the saying in language questions apart from a short period in the 50's and 60's when Greenlandic was somewhat stressed by Danish and Danish civil servants. With Home Rule in 1979 language again became an exclusive resort for the local Greenlandic authorities and substantial support was allocated to Greenlandic language and culture in a wider sense.

The situation is not very different in Canada and Alaska, especially in Canadian Inuktitut which in many respects is as well recognized and formally protected as is Greenlandic since the establishment of Nunavut 1999 and full transfer of political power in language questions to the local authorities in Iqaluit.

We are thus dealing with formally recognized languages in high esteem in their own societies and with access to quite substantial funding earmarked to projects in and for the local languages.

It ought to be comparatively easy to promote language technology in such a setting but the absence of not only language technology but also all kinds of basic language resources is striking.

And still more striking is the irresolution to get going even when offered the means to do so. During the International Polar Year three years ago a joint project between Nuuk, Tromsø and Odense offering language technology to Alaskan Iñupiaq and Canadian Inuktitut based on adaption of the Greenlandic automaton was unsuccessful mainly because local authorities in Alaska and Nunavut did not support the project. To mention one example.

The bottom line is thus a surprising mismatch between high language status and political attitudes in favour of language technology paired with funding possibilities that are not prohibitive on one hand and the fact that nothing happens on the other.

Why is it so difficult to get airborne?

There are, of course, many factors playing parts in the total explanatory framework but I would like to address one observation that as far as I know never has been treated in the literature

before, namely a skewness in Inuit languages' functions.

Until 1950 Greenland was monolingual with all parts of society carried out in Greenlandic. Mother tongue teaching worked according to the same scheme. The development of the subject progressed much like one would expect it to do in a modern society both didactically and technically.

In 1950 this state of affairs was dramatically altered when Greenland was decolonized. The Danish language attracted enormous status and very little attention was devoted to Greenlandic. As a consequence didactic development including production of teaching material for the mother tongue subject almost ceased.

By the mid 70-s it was a general belief that Greenlandic culture was moribund because of the pressure from Denmark and so was the culture's prime manifestation namely the Greenlandic language. The reaction was a culture revolution which ultimately paved the way for Home Rule with Greenlandic as the formal official language in 1979 and its expansion into Self Government in 2009.

Language questions played a significant role in the political movement in those days but the public debate about Greenlandic was thematically very different from the very vivid debate half a century earlier. It evolved around very general issues. Greenlandic was the prime ethnic, national symbol expressing Greenlandic culture and identity as an Inuk but very little energy was devoted to the instrumental and heuristic function of language as opposed to the debate before the war.

The same language view is found in mother tongue teaching in school and at the teacher training college leaving us with a whole generation of students and teachers with literally no descriptive framework for what is repeatedly stressed as one of the constituting parameters for their culture and personality.

To rephrase it we have a language in which the symbolic, artistic etc. half of the language by definition constitute all of the language leaving several functions of language entirely out of account.

Please observe, that I do not postulate such states of affairs to be inherent in Greenlandic and other minority languages for that matter. On the contrary do we know that Greenlandic philology and L1 didactics developed by native Greenlanders met very high standards before 1953 as long as Greenlandic developed at ease at its own pace. The skewness described above is a phenomenon that showed up after a period of pressure on Greenlandic.

The same phenomenon is evidently at play in a number of revitalized languages in Alaska and

Siberia so I believe we are dealing with a general process rather than a language specific one.

It should be obvious that the result inevitably will be a major conflict between language planning at the political level and corpus planning and modernization at the executive level. In a modern society quite a substantial bit of linguistic skills is needed to transform political decisions into everyday life applications. Love for one's language or artistic fluency in the language are important qualities but they do not compensate for lack of description.

Poor description creates poor teaching materials creates poor teaching creates poor motivation .. The descriptive incapability at all levels of society has given rise to a long row of problems and several nasty vicious spirals so that we have ended up in a deadlock situation with lots and lots of work to do but with no one to do it and much too limited tools to do anything really efficiently.

This is the exact reef where almost all projects with the least affinity to language technology are wrecked in spite of all the positive attitudes: Without precise descriptions of a language or staff with the skills needed to provide such descriptions technological solutions cannot be provided.

Problems to address and barriers to overcome

I think we all agree with the director-general of UNESCO when she states that technology is needed for the safeguarding and promotion of minority languages and linguistic diversity. The million dollar question is what we can allow ourselves to expect from inside the minority languages themselves.

Based on my experience expectations to language technology projects in minority languages must have an altered focus. It is next to impossible to find native speakers of for instance Greenlandic with the necessary ability to describe Greenlandic in terms concise enough for use in language technology projects. The approach to language description is simply entirely different - and, unfortunately, unusable with most language technology projects. And what is almost worse is the fact that I see no readiness among the elders who are the decision makers on indigenous culture and language to encourage alternative approaches to languages.

So to sum up: The first step in a language technology project for the next Inuit language in line for language technology is accordingly NOT to identify and define problems and

design a project that will deal efficiently with them. There are steps to take before that.

Establishing a language technology project for an Inuit language first of all depends on the elders' acceptance. In Canada and Alaska such acceptance must be formally obtained before establishing the project whereas it in Greenland is not a formal demand but rather an inevitable prerequisite for funding and access to resources and persons needed for the project since all questions concerning Greenlandic language will be passed on to the language board, Oqaasiliortut. Without Oqaasiliortut's approval projects do not have a chance in practice.

Once formalities are cleared and funding secured the question of locating manpower is next. With Canadian Inuktitut this has not at all been possible up till now, with Alaskan Iñupiaq in a joint project between Alaska Native Language Center and Carnegie Mellon a non-Inuk with a certain command of L2 Iñupiaq was hired to develop an automaton after many years of standstill because it wasn't possible to find the know-how needed for the project in Alaska. And in Greenland we have after years of serious problems with locating and retaining staff chosen to design the language technology program to include a formal education in language technology as a hands-on combination with the master of language technology program in Gothenburg.

Third web of problems arise from the lack of basic resources. Almost all linguistic resources at hand for Inuit languages are either rooted in the old missionaries' attempt to propagate Christianity or in attempts to translate foreign words and foreign concepts to the Inuit languages. Resources are accordingly almost exclusively bilingual or focusing on Inuit languages for L2 purposes whereas language internal resources like L1 grammar books or monolingual dictionaries and corpora are rudimentary or non-existing.

Finally, should an Inuit language technology project somehow overcome all the obstacles mentioned above the risk of drying out for lack of funding or drowning in success are both immediate because the public is incredibly attentive in language questions but expectations to technical solutions uttermost unrealistic.

Google Translate can be mentioned as one example. It is an often mentioned in the language debate in Greenland when critical voices rise from the political level as well as from the mediums. Instead of the - as it appears - unnecessary theoretical and tedious work with fst, CG and the like we could simply adapt Google Translate and other "off-the-shell" solutions.

Keep in mind that such opinions are aired by

persons in economic and political power but with basically no understanding of the language in descriptive terms. In that situation it is very hard to sit in the ivory tower and try to explain that data driven technology is not an option for a polysynthetic language and that we need endless years to pave the endless way via tagging and parsing toward rule-driven technologies.

Now, problems as the ones outlined here are of course all too well known to all of you. Still, in a micro first nation state like Greenland processes like these have immediate and direct impacts because we do not have the buffers of academic professionals in bureaucracy and universities to filter public opinions before they are taken to the political level. In Greenland we either have nothing at all between the public and the Parliament or we have institutions manned with lay people without theoretical schooling as is the case with Greenland's powerful Language Board. As a consequence we need to devote very much energy on "staying alive" that is to legitimize our project by answering scores of official memos and public reports, and by feeding the public with information about our doings with very small intervals.

Why free and ready accessibility is crucial to minority languages

Now, after all this lamentation you most likely have started to wonder what it all has to do with visibility and dissemination of language resources.

Very much, actually, so let me once again return to the opening of this presentation.

The Greenlandic project was established with very limited resources in terms of money, manpower, and know-how under the wings of the Sami project. It would obviously not be where it is now without the long-lasting support

from Tromsø. Greenland simply cannot itself provide the many tools needed and cannot maintain a forum strong enough to reinforce professionalism, pick up new trends and tools and secure transmission of skills to next generations.

The Sami project's definition of openness to include not only a download button one has to locate oneself but also a deliberate attempt to document and draw attention to resources paired with a willingness to invest time and energy in outsiders like myself paid off. It took quite some effort to launch the Greenlandic project but it functioned. We are still in business.

And it spreads as could be observed last year on Malta where an Iñupiaq project heavily inspired from the Greenlandic project was presented.

So to conclude this talk: Minority languages need language technology badly but very few have the human and linguistic resources needed to get going and the scantiest of resources namely the people in the projects will inevitably find themselves spending most of their time not on language and technology but on human resource development, bureaucracy, and public promotion just to keep a project alive.

The bottom line then is twofold:

(i) We need help, and lots of help at that. Therefore easy access not only to resources but also to actual programs and tools is not only welcomed but rather the very lifeline for a project like the Greenlandic project. We are much too few and we still need all kinds of resources so our only option is to borrow or steal whatever can be borrowed or stolen and limit local resources to deal with language specific and culture specific problems that under no circumstances can be outsourced.

(ii) It works. The Greenlandic language technology program has proven that it can be done in spite of everything when good forces are pooled consciously and deliberately.