

Liv Molich

- opnåede titlen Academicus Arcticus og vandt tredjepræmien på 2.500 kr. ved prisuddelingen 2. november på Arktisk Institut i København.

Liv Molich er kandidatstuderende på Ilisimatusarfik, Grønlands Universitet og ansat ved Oqaasileriffik, Grønlands Sprogsekretariat.



Kan en computer oversætte mellem grønlandsk og dansk?

Grønlandsk og dansk er temmelig forskellige sprog, og det kan derfor være svært at oversætte mellem dem. Heldigvis er der efterhånden udviklet en hel del computerbaserede hjælpemidler, som gør det nemmere. Hos Oqaasileriffik (Grønlands Sprogsekretariat) sidder en lille gruppe medarbejdere og arbejder med sprogteknologiske projekter. Sidste skud på stammen er en oversættelsesmaskine, som er under udvikling.

To forskellige sprog

Det grønlandske og det danske sprog har hver sin historie og oprindelse. Grønlandsk er et af de såkaldte eskimoisk-aleutiske sprog, der tales i de kystnære områder rundt om Polarcirklen, fra Grønland i øst, hen over Canada og Alaska og helt hen til Sibirien lige vest for Beringstrædet. Dansk hører derimod til den indoeuropæiske sprogfamilie, der også omfatter blandt andet engelsk, latin, græsk og hindi.

Hvis man sammenligner sprogene, er især én forskel iøjnespringende: Grønlandsk er et polysyntetisk sprog, hvor ét ord ofte udgør en hel sætning. Mange af de grønlandske ord er derfor lange og unikke, og netop fordi de er unikke, kan de ikke uden videre slås op i en ordbog. Derimod har dansk mange små ord, som forekommer ofte, og som står i ordbogen. Sprogenes forskellige struktur gør, at maskinoversættelse indebærer en hel masse analysearbejde og ordbogsarbejde.

```
LUUNNIIT_PART (L) :eller; P-1=(<Conj-encl>) :eller_[@ii->@CO]
H=("suna" @OBJ)_noget :som=helst; P-1=("kina|suna") :som=hel
num)>) :end (end ikke en eneste); H=(CON) P-1=(<Adv-encl>) :se
>@SUB] [PART->KS] P-1=('GALUAR')_nil ('uanset' vil i nogle tilf
eneste mulighed, men 'selvom' som regele bedre: Piumasaraara p
sorliugaluarpalluunniit iliuuserisimasaminut akisussaatinneqa
(H=(V CON) B=(V CON) :eller_[@ii->@CO] [PART->KC]; S=(@SUB) H=(
><noks>) :selvom_ [PART->KS] P-1=('GALUAR')_nil; S!=(@CO|@SUB)
PROP) :selv_ [PART->ADV] [.*->@>N], :mere=end=selv (?); D!=(@FO
H=(@PRED)_[<ø>-><noks>] :eller_[@ii->@CO] [PART->KC] P-1=('GALU
=(@FOC>) S!=(@SUB) H=(@ii->V)_[<ø>-><noks>] GM=(@PRED) :eller_
[PART->KC] P-1=('GALUAR')_nil; D!=(@FOC>) S!=(@CO) H=(V)_[<ø>-
><noks>) :uanset_ [PART->SUB] [PART->KS] P-1=('GALUAR')_nil; H=(
encl>) :endog; B=(<Adv-encl>) :uanset, :endog; DB
```

Figur 1: Uddrag af den ordbog, der bliver brugt til maskinoversættelse. Reglerne beskriver, hvornår -luunniit skal oversættes til "eller", "som helst", "end", "selvom", "selv", "uanset" og "endog".

– men med ligheder

To af de vigtigste sprogvidenskabelige discipliner kaldes morfologi og syntaks. Her studerer man, hvad de enkelte ord består af, og hvordan ordene danner sætninger. Resultaterne kan man læse i ordbøger, grammatikker og videnskabelige afhandlinger.

Der findes grønlandske ordlister helt tilbage fra 1500- og 1600-tallet, og fra 1721 udvikles en egentlig skriftkultur i Grønland, i begyndelsen med missionær Hans Egede i spidsen. Hans ordlister


bliver startskuddet til en stribe af grønlandske ordlister, ordbøger, grammatikker og lærebøger. Vi har altså 300 års erfaring med grønlandsk sprogbeskrivelse.


Ved at studere sprogbeskrivelserne finder man hurtigt ud af, at mange grønlandske ord kan deles op i mindre bidder, der hver for sig svarer til hele eller dele af danske ord. Det danner basis for maskinoversættelse – alle sprog har ligheder, og det er disse ligheder, vi udnytter, når vi oversætter!


Den digitale tidsalder

I dag findes flere af de grønlandske ordbøger og grammatikker digitalt, og med den digitale tidsalder er der også blevet udviklet adskillige andre computerbaserede hjælpemidler til grønlandsk: Stavekontrollen *Kukkuniiat* sørger for, at forkert stavede ord bliver markeret, analyseværktøjerne *Ordanalysator*, *GroenOrd*, *LG3* og *Liveanalyse* giver analyse og oversættelse af ord og sætninger, talesyntesen *Martha* kan læse teksten op, og onlineordbogen *Katersat* er for længst blevet større end alle tidligere ordbøger, selvom den langt fra er færdig endnu.

Ordbog

oqaasilerineq t. (KAL.) 

Relativ kasus: oqaasilerinerup 

Pluralis: oqaasilerinerit 

oqaaseq+LIRI+Der/nv+NIQ+Der/vn+N+Rel+Sg

inuaat oqaasiisa atornerinut malittarisassat; soorlu kalaallisut immikkut oqaasileruteqartoq; takukkit: uiguutilerineq aamma oqaaseqatigiilerineq

Relaterede ord: [grammatik](#) ; **Dansk:** [grammatik](#) ; **Engelsk:** [grammar](#)

Brugseksempler: [grammatikki](#); [oqaasilerissutit](#) ; **Dansk:** [en grammatisk analyse](#)

Referencer: [Erik F, Oqaatsit 1990, Taaguusersuutit - kalaallisut qallunaatullu atuartitsinermi](#)

Figur 2: Eksempel på opslag i onlineordbogen Katersat. Her kan man få ordet analyseret, læst højt, forklaret og oversat.

En del af disse ressourcer er udviklet af Oqaasileriffik på baggrund af de ældre og nyere sprogbeskrivelser, og de udgør grundlaget for maskinoversættelse. Maskinoversættelse kan sammenlignes med Google Translate; men Google Translate eksisterer ikke for grønlandsk og vil aldrig kunne komme til det, fordi dens teknologi baserer sig på statistik. Det kræver blandt andet, at det samme ord forekommer mange gange, og det er ikke tilfældet for de lange grønlandske ord. Derfor kan rent statistisk oversættelse ikke bruges til grønlandsk. Derimod kan man bruge såkaldt regelbaseret oversættelse.

Maskinoversættelse bygger bro

Med regelbaseret teknologi er det muligt at udvikle maskinoversættelse mellem grønlandsk og dansk. Maskinoversættelser kan ikke bruges til at lære et sprog med, men de kan lette

kommunikationen mellem to mennesker, der ikke taler hinandens sprog. Som chefkonsulent ved Oqaasileriffik Per Langgård udtalte allerede i 2002: “Grønland har behov for grønlandssproget teknologi af praktisk kommunikative og økonomiske årsager, simpelthen fordi indbyggerne i Grønland taler grønlandsk (...)”

Maskinoversættelsesprojektet *Nutserut*, som er et projekt under Oqaasileriffik i samarbejde med det danske firma GrammarSoft, forbedrer og udnytter en række sprogteknologiske værktøjer, som allerede længe har været at finde på oqaasileriffik.gl. De bliver kombineret til en maskine, der i løbet af en indviklet proces oversætter ord og sætninger fra dansk til grønlandsk og fra grønlandsk til dansk.

Qarasaasiami programmi atorlugu
tamatigoornerusunik nuilarmiuliortoqarsinnaavoq

Qarasaasiami programmi atorlugu
tamatigoor NIRU TUQmik nuilarmiut LIUR TUQ QAR SINNAavoq

Der SINNAA atorlugu qarasaasiami programmi
QAR TUQ LIUR en nuilarmiut TUQ tamatigoor mere alsidig

Der kan ved brug af computerprogrammet
være nogen som laver en perlekrave som er mere alsidige

Figur 3: Oversættelse fra grønlandsk (grøn) til dansk (rød). Den grønlandske sætning bliver først brudt op i mindre dele. Dernæst bliver delene flyttet, så sætningen får dansk ordstilling, og enkelte danske ord bliver tilføjet. Til sidst bliver det hele sat sammen til en sætning og oversat. Illustrationen viser en simplificering af processen. Den grønlandske sætning er taget fra sermitsiaq.ag.

Arbejdet er ikke bragt til ende

Eksemplet i figur 3 illustrerer, hvordan oversættelsesmaskinen håndterer oversættelse fra grønlandsk til dansk, men viser også, at oversættelsen kan blive bedre. Sermitsiaqs danske version af teksten, “Computerprogram sikrer mere mangfoldige perlekraver”, er mere elegant end maskinoversættelsen, men alligevel er det tydeligt, hvad den maskinoversatte sætning betyder – og resultatet er langt bedre, end hvad de fleste dansktalende eller Google Translate formår.

Moderne teknologi fremtidssikrer 300 års grønlandsk sprogbeskrivelse og gør os fortsat klogere på sprogets mange finurligheder. Vi har en maskine, der kan oversætte mellem grønlandsk og dansk, og dag for dag gør *Nutserut*-medarbejderne den bedre og bedre; men ligesom Google Translate ikke oversætter perfekt, vil den grønlandske oversættelsesmaskine aldrig kunne erstatte rigtige tolke og oversættere.

Referencer

- Aagesen, Henrik Vagn (2008): *Grønlændernes Ord*. <http://www.groenord.dk/>.
- Berthelsen, Chr. et al. (2006): *Oqaatsit. Kalaallisuumiit Qallunaatumut* (ordbog). Nuuk: Ilinniuserfik.
- Jacobsen, Birgitte (1997): "OQAATSIT i tid og rum. Historie og status." *LexicoNordica* 4, 41-64.
- Jones, Kirsten Gade og Robert Petersen (2012): *Oqaasersiuut* (ordbog). Nuuk: Ilinniuserfik.
- Langgård, Per (2002): *Sprogteknologi i Grønland*. Rapport til Nordisk Forskeruddannelsesakademi præsenteret i juni 2002, <https://oqaasileriffik.gl/udgivelser/om-groenlandsk-sprog/>.
- Langgård, Per (2015): "Selvfølgelig snakker vaskemaskiner grønlandsk i fremtiden." *Språk i Norden* 2014, 110-125, <https://tidsskrift.dk/sin/article/view/20185>.
- LearnGreenlandic (2017): *LG3*, <https://www.learngreenlandic.com/lg3/>.
- Nielsen, Flemming A. J. (2011): "Den ældste grønlandske Bibel – et sprogligt og kulturelt møde." I: Høiris, Ole og Ole Marquardt (red.): *Fra vild til verdensborger*, 147-190. Århus: Aarhus Universitetsforlag.
- Oqaasileriffik: Sprogteknologiske hjælpemidler, bl.a. *Katersat* (ordbog), *Kukkuniii* (stavekontrol), *Liveanalyse* (sætningsanalyse), *Martha* (talesyntese) og *Ordanalysator*, <https://www.oqaasileriffik.gl/>.
- Oqaasileriffik (2010): *A bit of history*, <https://oqaasileriffik.gl/langtech/a-bit-of-history/>.
- Oqaasileriffik (2016): *Ukiuimoortumik nalunaarut 2016-imoortoq* [årsrapport 2016], <https://oqaasileriffik.gl/wp-content/uploads/2018/08/2016-Oqaasileriffik-final.pdf>.