

Slutrapport språkteknologisk pilotsatsning på de små språken i Norden för att nå slutanvändare

Under 2022 och 2023 utfördes en språkteknologisk utbildningsatsning på de små språken i Norden. Målet var att sprida kunskap om vilka verktyg och resurser som finns för de små språken. Målgruppen för satsningen var slutanvändare och för att nå dem utbildade vi först sex olika språkambassadörer som sedan fick sprida sina kunskaper vidare i sina respektive språksamhällen.

Bakgrund

Den här satsningen grundar sig i det nordiska initiativet som Ad hoc-gruppen för Små språk i Norden har initierat för språkteknologi som syftar till att stärka den digitala inkluderingen (se bilaga 1 för hela det nordiska initiativet). Initiativet består av tre delinitiativ:

- **Initiativ 1:** 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.
- **Initiativ 2:** Gemensam nordisk hållning och samarbete för att inkludera de små språken i EU:s arbete med språklig jämlikhet i Europa 2030.
- **Initiativ 3:** Gemensam nordisk dialog med och strategi för att driva på tech-giganternas arbete med och inkludering av små språk i Norden.

Den här satsningen kan knytas an till det första delinitiativet.

Språkteknologi och de små språken i Norden

För flera av de små och mer resurssvaga språken ligger utvecklingen av språkteknologi långt efter i jämförelse med de större språken i Norden och övriga världen, exempelvis i fråga om tillgången till rättstavningsprogram, program för språkinlärning och digitala ordböcker. Det finns en stor enighet om att utvecklad och spridd språkteknologi är en förutsättning för att stärka och revitalisera små språk.

Dagens användare av de små språken är lika beroende av att kunna använda mobiltelefoner, datorer och andra digitala verktyg på sina språk som de som talar och använder större språk. Norge, via bland annat Tromsø universitet, har utvecklat åtskilliga språkteknologiska verktyg. Kunskapen om dessa språkteknologiska lösningar och verktyg och hur de kan användas är dock låg hos utbildningssektorn, språkanvändare och myndigheter som ska använda språken i de nordiska länderna och i de självstyrande områdena.

Satsningens mål

För att tillgängliggöra redan befintlig språkteknologi för slutanvändare behövs ökad kunskap om de senaste språkteknologiska verktygen samt praktisk hjälp för hur dessa används. Därför har denna satsning fokuserat på att öka kunskapen hos slutanvändarna så att de kan använda de språkteknologiska verktygen för sin egen eller andras språkanvändning och språkinläring. Målgruppen slutanvändare är bland annat modersmåls- och språklärare, studenter och elever, handläggare och andra som har behov av verktygen.

Så här kan man sammanfatta målen med satsningen:

- Utbilda språkambassadörer som arbetar med små, nordiska språk.
- Skapa en miljö där språkambassadörerna kan lära av varandra och lära av varandras erfarenheter.
- Nå ut med kunskapen om språkteknologi till slutanvändare.

Deltagare och språk

I satsningen ingick sex språk: nordsamiska, sydsamiska, lulesamiska, kvänska, meänkieli och färöiska. Gemensamt för språken är att det finns grundläggande språkteknologisk infrastruktur som kan läras ut till slutanvändare. Trots det har de olika språken väldigt olika mycket tillgång till språkteknologiska verktyg. Till exempel på nordsamiska finns avancerade verktyg som maskinöversättning, medan på meänkieli finns endast tillgång till en handfull mindre avancerade verktyg, såsom en digital ordbok och korpus.

För varje språk nämnt ovan deltog en så kallad ”språkambassadör”. Utöver språkambassadörerna deltog Grönlands språksekretariat, Oqasileriffik, som gott exempel och Trond Trosterud från Giellatekno, Universitetet i Tromsø deltog som ansvarig för den språkteknologiska utbildningen i aktivitet ett.

Språkambassadörerna för respektive språk:

- Nordsamiska, Marit Elisabeth Eira som är lärare i nordsamiska, Norge
- Sydsamiska, Anders Östergren Njajta, sydsamisk språkkonsulent på Sametinget, Sverige.
- Lulesamiska, Jörn-Robert Knutsen, språkmedarbetare på Årran lulesamiskt språkcentrum, Norge.
- Kvänska, Mari Keränen, Kvenskt institutt, Norge.
- Meänkieli, Elina Kangas, Institutet för språk och folkminnen, Sverige.
- Färöiska, Uni Johannesen, masterstudent i färöiska på Färöarnas universitet, Färöarna.

Utöver dessa officiella deltagare hade vi ytterligare deltagare på kurserna som projektet inte betalade för eftersom intresset från språkarbetare var stort. Det var två som arbetade med kvänska: Aili Eriksen från Kvensk institutt och Konstantina Vardaka från Storfjord språkcentrum. Det var även

två som arbetade med nordsamiska: Anna Wasara och Reetta Tronensis från svenska Sametingets nya språkcentrum i Kiruna och Jokkmokk. Extradeltagarna hade förstås inget krav på sig att genomföra några aktiviteter, utan deltog av eget intresse.

Aktiviteter

I satsningen genomfördes tre olika aktiviteter. De första två handlade om att utbilda språkambassadörerna och den sista gick ut på att de fick hålla egna workshoppar eller liknande för att nå ut till slutanvändare.

Aktivitet 1 – Kurs i installation och användning av språkvetenskapliga program för nordiska minoritetsspråk vid Tromsø universitet

Satsningens första aktivitet var en fem dagars utbildning hos Giellatekno vid Universitetet i Tromsø, UiT. Giellatekno, tillsammans med Divvun-gruppen. På UiT har utvecklats åtskilliga språkvetenskapliga verktyg för nordiska minoritetsspråk. De arbetar mycket med uraliska språk, såsom samiska och kvänska, men även andra nordiska minoritetsspråk och små språk. Trond Trosterud, professor i samiska, höll i utbildningen.

Utbildningen pågick under fem dagar, 29 augusti till 2 september 2022. Under utbildningen gick vi bland annat igenom hur grammatiska modeller i rättstavningsprogram fungerar, hur man installerar tangentbord för olika operativsystem, hur man använder översättningsminnen och hur nutiden och framtiden ser ut för ljudteknologi (talsyntes och taligenkänning) för små språk.

Kursansvarige Trond Trosterud gick dels igenom generella, språkvetenskapliga problem och hjälpte deltagarna installera de språkvetenskapliga verktyg som finns. Målet med den här aktiviteten var dels att deltagarna skulle lära sig om hur språkvetenskap kan stärka språken de arbetar med, och hur de kan lära slutanvändare att installera och använda de språkvetenskapliga verktygen.

Förutom att kursdeltagarna fick lära sig om verktygen uppstod diskussioner om framtida behov mellan kursdeltagarna och Giellatekno och Divvun. Dessutom kunde eventuella tekniska problem upptäckas och lösas på plats när många olika använde verktygen.

På kursen närvarade totalt tio deltagare: fem av språkambassadörerna (nordsamiska, lulesamiska, kvänska, meänkieli, färöiska), fyra övriga språkarbetare och projektsamordnaren.

Aktivitet 2 – Workshop om språkvetenskap med Grönlands språksekretariat Oqaasileriffik

Del två av projektet bestod av en workshop som Grönlands språksekretariat, Oqaasileriffik anordnade i Köpenhamn. Från Grönlands språksekretariat var det Per Langgård, Beatrine Heilmann och Judithe Denbæk som berättade om deras arbete med språkvetenskap. Anledningen till att Grönlands språksekretariat användes som gott exempel är att de kommit väldigt långt i arbetet med språkvetenskap.

På workshoppen i Köpenhamn fick deltagarna bland annat lära sig om arbetet med den grönländska stavningsskontrollen och språkmodellen, hur normeringen av det grönländska skriftspråket var en viktig förutsättning för arbetet med språkvetenskap och hur de tagit fram andra språkvetenskapliga verktyg, som maskinöversättning.

Egna workshoppar

Den sista och viktigaste delen av projektet var förstås att nå slutanvändare, vilket kunde göras på en rad olika sätt. Eftersom deltagarna har väldigt god insyn i sina språksamhällen kunde de anpassa den här aktiviteten efter vad de ansåg skulle passa bäst utifrån språksamhället. Dessutom skilde sig tillgången till språkteknologiska verktyg mycket mellan språken, vilket också påverkade hur deltagarna valde att nå ut till slutanvändarna. Det kommer jag gå in på i mer detalj när jag diskuterar respektive språk.

Tyvärr blev det ingen workshop för sydsamiska i det här projektet eftersom de inte hann med det. Men förhoppningsvis kommer erfarenheterna från det här projektet ändå komma till nytta längre fram.

Nordsamiska

För det nordsamiska språksamhället höll Marit Elisabeth Eira en presentation på Nord-Troms vidaregående skola för lärare i samiska språk. I programmet ingick följande:

- bakgrund till utbildningssatsningen och varför språkteknologi är viktigt för små språk
- vilka verktyg som finns på nordsamiska
- hur man installerar verktygen på PC
- hur man installerar verktygen på Mac
- hur man installerar verktygen på mobil.

Utöver det delade hon med sig av andra resurser, såsom Divvuns och Giellateknos hemsida och [Youtubekanal](#) där det finns ytterligare instruktioner.

Lulesamiska

Jörn-Robert Knutsen från Årran lulesamiskt språkcentrum höll också en presentation om de vilka språkresurser som finns på Lulesamiska i samband med den samiska språkveckan 2022. Bland annat det här gick han igenom under sin presentation:

- bakgrund till utbildningssatsningen och varför språkteknologi är viktigt för små språk
- hur man använder de verktyg som finns online, till exempel Korp, digitala ordböcker och ordgenerator
- hur man installerar och använder stavningskontrollen.

Förutom att presentationen nådde de som var på plats laddades det även upp på [Youtube](#).

Nu planerar Årran språkcentrum även att göra instruktionsvideor på Youtube och ladda upp det på hemsidan [arran.no](#). De har även lagt länkar till Giellatekno och Divvun under ”[språkresurser](#)”.

Kvänska

Mari Keränen från Kvensk institutt höll en kombinerad skriv- och språkteknologikurs som riktade sig mot skribenter som skriver på kvänska. Här är ett utdrag från inbjudan:

Kainun institutti – Kvensk institutt arrangerer et kombinert språk- og skrivekurs 15.–17.2.2023 i Børselv.

Vi starter med et halvdags språkkurs 15.2. etter lunsj, fortsetter med introduksjon til språkteknologiske verktøy 16.2. og avslutter med et halvdags skrivekurs 17.2. Det er mulig å delta på en, to eller alle tre dagene.

Målgruppen er de som skriver på kvensk i sin jobb, dvs. journalister, lærere, freelance-oversettere, formidlere osv., men hvis vi har ledige plasser så åpner vi til tilbudet til alle interesserte. (Kvensk institutt, <https://www.kvenskinstitutt.no/2022/11/16/sprak-og-skrivekurs-15-17-2-2023-i-borselv/>)

I den språkteknologiska delen av kursen ingick bland annat följande:

- Vilka språkteknologiska verktyg som finns för kvänska.
- Hur man installerar och använder dem.
- Hur man använder de språkteknologiska verktygen när man skriver text.

Kursen var lyckad och Mari som höll i den fick mycket positiv feedback av de 15 som deltog. Dessutom har andra kvänkspråkiga institutioner visat intresse för att beställa liknande kurser, vilket tyder på ett växande intresse inom språksamhället.

Meänkieli

Elina Kangas höll i ett hybridseminarium om språkteknologi och språkvård på meänkieli i Stockholm. Hon höll det tillsammans med Meri Alarcón, som också är språkvårdare i meänkieli på Isof, och Rickard Domeij och Marie Mattson som båda arbetar med språkteknologi på Isof. På seminariet gick hon både igenom bakgrunden till språkteknologi och varför det är viktigt samt, hur man just de verktyg som finns på meänkieli, såsom korpus och den digitala ordboken. Utöver det diskuterade hon framtiden för språkteknologi på meänkieli, bland annat att Isof just nu arbetar på att ta fram en stavningskontroll.

Färöiska

Uni Johannessen höll flera workshoppar på sex olika skolor runtom Färöarna (2 i Klaksvík, 1 i Vestmanna, 1 på Kambsdalur, 1 i Tórshavn och 1 på Suðuroy). Det här gick han igenom:

- Varför språkteknologi är viktigt för det färöiska språket.
- Hur man installerar och använder de språkteknologiska verktygen.

Det var cirka 5 – 15 deltagare på varje workshop, vilket innebär att projektet nådde cirka 60 personer som arbetar med färöiska.

Övriga resultat av satsningen

Ett av de viktigaste resultaten har varit att diskursen om språkteknologi för små språk har varit mer aktiv. Bland annat har NRK kvensk skrivit om satsningen (<https://www.nrk.no/kvensk/unike-sprakverktoy-brukes-for-a-styrke-sma-nordiske-sprak-1.16094738>). Även Henrik Niva, Maria Heikka och Elina Kangas pratade om satsningen på meänraatio 9 augsti (inspelningen är nu nedtagen).

Sammanfattning projektrapport

Under hösten och vintern 2022 – 2023 har Isof samordnat en språkteknologisk utbildningssatsning för att nå ut med information om språkteknologiska verktyg till slutanvändare. För att nå slutanvändare utbildades språkambassadörer i sex olika språk vid Universitet i Tromsø som arbetar mycket med språkteknologi för små språk. De fick även delta på en workshop som handlade om hur Grönlands språksekretariat arbetat med språkteknologi, där de användes för gott exempel.

Projektet resulterade i att fem av sex deltagare höll i olika typer av lyckade utbildningssatsningar i sina språksamhällen. Dessutom har projektet resulterat i att flera språkarbetare i nordiska länder har fått möjlighet att diskutera den här typen av frågor med människor som arbetar med språk i liknande situationer.

Intresset för projektet har varit stort. Dels kom det fler deltagare på utbildningen än vad som var planerat för, och dels har flera medier skrivit om satsningen.

Språkteknologisk utbildningssatsning för små språk i Norden
Marie Mattson, Institutet för Språk och folkminnen
Rapporteringsdatum 2023-03-30
Projektperiod 2019-04-15- 2022-12-31

Bilaga 1

Nordiskt initiativ för språkteknologi – digital inkludering av små språk i Norden

Sammanfattning

Det saknas i stora delar tillgång till grundläggande samhällsviktig infrastruktur som möjliggör digital kommunikation på de små språken¹ i Norden. Den digitala exkluderingen påverkar hundratusentals människor i den nordiska gemenskapen, detta trots att språken har officiell status i de nordiska länderna. Den digitala exkluderingen omfattar alla de små språken i Norden i större eller mindre omfattning.

Nordiska ministerrådet uppmanas att genomföra ett samlat initiativ i tre delar inom det språkteknologiska området i syfte att stärka de små språken i Norden.

Språkteknologi bidrar till en stärkt nordisk gemenskap

Alltmer av vårt samhällsdeltagande sker via internet. Där får vi viktig information och kan kommunicera med myndigheter, familj och vänner. Vi streamar nyheter och använder digitala tjänster. Vi tänker sällan på att det krävs avancerad teknik för att använda våra nordiska språk i den digitala världen. Språkteknologi är som en osynlig hand som hjälper oss att hitta information, kommunicera och lära oss.

För många små språk i Norden saknas språkteknologi, den är ofullständig eller kan inte användas. Saker som borde vara självklara – så som att skriva och använda språket i digitala sammanhang – är svåra att göra på små språk så som till exempel de samiska språken, kvenska, färöiska, meänkieli, romani och romanes.

Det betyder att många nordbor saknar tillgång till den grundläggande, samhällsviktiga infrastruktur som språkteknologin garanterar. Så är fallet samtidigt som länderna har åtagit sig att skydda och stärka de små språken².

Tillsammans har de nordiska länderna och särskilda områdena kraft och resurs till att genomföra ett gemensamt initiativ för att garantera användare av små språk en inkludering i den digitala språkliga gemenskapen och därigenom stärka den nordiska sammanhållningen. För att möjliggöra detta uppmanas Nordiska ministerrådet att genomföra tre initiativ för språkteknologi.

Initiativ 1: 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.

Den samlade nordiska hållningen och de enskilda nordiska staternas skyldigheter³ är att de nordbor som använder de små språken ges möjligheter att lära sig, använda och vidareutveckla språken och att utöva samt utveckla den egna kulturella identiteten.

De nordiska länderna och särskilda områdena har en stor möjlighet att inom det språkteknologiska

¹ Med små språk menas i det här sammanhanget små språk i Norden som har relativt få talare med officiell status, inklusive urfolkspråk, nationella minoritetsspråk och teckenspråk.

² Deklaration om nordisk språkpolitik

³ Exempelvis teckenspråkslagen (359/2015) i Finland, språklöva (Lovvedtak 89, 2020–2021) i Norge och Lög um stöðu íslenskrar tungu og íslensks táknaðs (Lög nr. 61 7. júní 2011) i Island.

området lära av varandra och dra nytta av de erfarenheter som finns inom Norden av funktionell språkteknologi och genom samverkan bidra till förbättrad digital inkludering. I det språkteknologiska utvecklingsarbetet bör även ingå att kartlägga resurser och behov för de små språken i Norden som grund för kommande strategiska satsningar.

Initiativ 2: Gemensam nordisk hållning och samarbete för att inkludera de små språken i EU:s arbete med språklig jämlikhet i Europa 2030.

EU satsar stora resurser på att utveckla språkteknologi för att göra ett flerspråkigt Europa möjligt. EU-gemenskapen låter utveckla avancerade maskinöversättningssystem som ska kunna användas för att översätta information och tjänster över språk- och nationsgränserna. Det uttalade EU-målet är att uppnå digital språklig jämlikhet i Europa 2030⁴.

Den nordiska gemenskapen kan genom ett initiativ för digital inkludering av små språk bli föregångsländer genom att lyfta in dessa i EU:s arbete för språklig jämlikhet i Europa och uppmärksamma deras behov av språkteknologiska lösningar. Hittills har EU-arbetet nästan uteslutande fokuserat på de officiella språken inom EU.

Initiativ 3: Gemensam nordisk dialog med och strategi för att driva på tech-giganternas arbete med och inkludering av små språk i Norden.

Med digitaliseringen har mer och mer av inflytandet – och därmed makten – över språken hamnat hos de stora tech-giganterna. Multinationella bolag äger de digitala plattformar och verktyg som är en del av de nordiska medborgarnas vardag.

Tech-bolagens makt över våra språk är stor och de har inflytandet att styra över de små språkens framtid. Erfarenheter från inte minst Island, Norge och Grönland visar samtidigt att det för ett enskilt land eller en språkgrupp är komplicerat att föra konstruktiva dialoger med tech-giganterna.

Genom ett gemensamt språkpolitiskt agerande skulle de nordiska länderna och de särskilda områdena kunna utgöra en numerärt och betydelsemässigt relevant partner för tech-bolagen. Ett sådant samarbete skulle vara bra för såväl tech-bolagen som de nordiska länderna och särskilda områdena och inte minst de nordiska medborgarna.

Stöd för och undertecknande

Det här nordiska initiativet för språkteknologi – digital tillgänglighet och inkludering av små språk i Norden är initierat och stöds av ad hoc-gruppen för Nordiska ministerrådets insats Små språk i Norden (2019–2022) och utarbetat av och undertecknad av sakkunniga inom området. I ad hoc-gruppen ingår representanter för de nordiska länderna och särskilda områdena.

Initiativet stöds vidare av European Federation of National Institutions for Language (EFNIL) och Permanent International Committee of Linguists (CIPL) samt Arbetsgruppen för språkteknologi i Norden (Astin).

⁴ <https://european-language-equality.eu/>

2023-03-30

Handläggare
Marie Mattson

Kunnskapsdepartementet/Ministry of
Research and Education

Rickard Domeji
Ansvarig för språkteknologi, Institutet för språk och folkminnen, Sverige

Sólvá Jónsdóttir
Sakkunnig färöisk, Nám, Färöarna

Sjur Nørstebo Moshagen
Chefsingenjör Divvun-gruppa, UiT Noregs arktiske universitet

Per Langgård
Chefkonsulent Oqaasileriffik, Grönland

Trond Trosterud
Professor i samisk språkteknologi, UiT Noregs arktiske universitet