

Språkvård och språkteknologi i Norden

Av Rickard Domeij og Torbjørg Breivik

Ett helt avgörande instrument för människors möjligheter att använda modern informations- och kommunikationsteknologi är språket. Finns tekniken och tjänsterna tillgängliga på det egna språket, har man större möjligheter att tillgodogöra sig dem. Språknämnderna i Norden har i olika grad uppmärksammat och ägnat sig åt det språkliga innehållet hos den nya tekniken. Artikeln ger en översikt över vad språknämnderna gör och har gjort var för sig, och vilka initiativ till samarbete och samordning som finns.

På senare år har språkteknologi blivit ett uppmärksammat område i de nordiska ländernas språkvårdsarbete. Man har insett att språkvården inte klarar sig utan språkteknologi, men också att språkvården kan spela en viktig roll för språkteknologin, inte minst genom att bidra till att förbättra den språkliga kvaliteten hos språkteknologiska produkter. Men varje nordisk språknämnd för sig har inte resurser och kompetens att utveckla det språkteknologiska arbetet som behövs. Därför finns mycket att vinna på att samla de nordiska ländernas språkteknologiska kompetens i en gemensam resurs.

Fördelarna med ett nordiskt samarbete om språkteknologi diskuterades på ett nordiskt seminarium i Pargas våren 2005. Som en följd av seminariet beslutade Nordens språkråd att låta utreda möjligheterna till en gemensam nordisk satsning på språkteknologiområdet i en vismansrapport som ska vara klar våren 2006. Det bildades också en arbetsgrupp med representanter från språknämnderna i Norden med syfte att få igång ett samarbete om språkteknologi.

Samarbete om språkteknologi

Redan i dag kan man se exempel på samarbete mellan språkvård och språkteknologi i de enskilda länderna. Ett exempel är arbetet med den norska handlingsplanen för språk och informationsteknologi, vilken är en del av eNorge-planen. Handlingsplanen och den uppföljande projektplanen för insamling och tillgängliggörande av språkteknologiska resurser (norsk språkbank) är resultatet av ett samarbete mellan Språkrådet och språkteknologisamhället i Norge (både forskning och industri).

Nödvändigheten av en norsk språkbank ligger i det faktum att marknaden för norskspråkiga produkter är begränsad, och att det är dyrt att samla in och göra språkliga resurser tillgängliga för språkteknologiska ändamål. Tanken är att språkbanken ska vara en gemensam samhällelig resurs som finansieras med offentliga medel och ska ställas till förfogande för forskning och utveckling. Ser man till hela den nordiska marknaden, som inte heller den är särskilt stor i sammanhanget, finns också där mycket att vinna på att göra gemensam sak när det gäller resurssamlingar och språkteknologisk kompetens, som är en bristvara. Men låt oss börja med att se vilka språkteknologiskt relaterade aktiviteter som pågår i anslutning till språkvårdsarbetet i Norden.

Svenska språknämnden, Sverige

Svenska språknämnden ägnar en viss del av sin verksamhet åt språkteknologi. Det gäller framför allt skrivstöd som stavnings- och grammatikkontroll, som har en språknormerande roll. Språknämnden har deltagit i diverse forskningsprojekt som, förutom forskningsrapporter (se t.ex. Karlsson, 2004), resulterat i språkgranskningsprogrammet Granska och den webbaserade språkinlärningsmiljön Grim¹. Förra året gav språknämnden ut boken Datorn

¹ <http://skrutten.nada.kth.se/grim/>

granskar språket (Domeij, 2005) som ger en populär beskrivning av språkkontrollens funktioner och användbarhet. Nyligen har innehållet i Words grammatikkontroll granskats och kommenterats från språkriktighetssynpunkt i samarbete med programutvecklaren (Karlsson, 2005). Språknämnden är också starkt delaktig i driften av Språkteknologi.se², en webbplats med information om svensk språkteknologi.

När Svenska språknämnden ombildas till myndighet den 1 juli 2006 får den ett större ansvar för det språkteknologiska området som helhet. I propositionen *Bästa språket* (2005) sägs att den nya myndigheten ska ha en ”samordnande språkteknologisk funktion” med syfte att stärka de statliga insatserna på språkteknologi. Myndigheten ska inte bara följa utvecklingen på området och uppmärksamma teknikens effekter på språk och skrivande; den ska också peka på brister i tekniken eller på avsaknad av svenskspråkig teknik och agera för att åtgärda dessa brister på olika sätt.

Förhoppningsvis kan myndigheten bli en språkpolitisk plattform till gagn för forskning och utveckling av god språkteknologi, inte bara för svenska utan också med hänsyn till minoritetsspråken, invandrarpråken, teckenspråket och andra särskilda behovsområden. Problemet är att möjligheterna till insatser är begränsade eftersom resurserna är små. Med ett väl fungerande nordiskt samarbete blir möjligheterna större.

Språkrådet, Norge

Sedan år 2000 har Språkrådet i Norge två personer som arbetar med det norska språket och informations- och kommunikationsteknologin (IKT). Det kan ses som ett uttryck för uppmärksammandet av ny teknologi och omläggningen till digital ärendebehandling, självbetjäning via internet vid offentliga tjänster och diskussioner om hur alla skulle få tillgång till den nya tekniken och kompetens att använda den. Sådana frågor stod på den norska regeringens dagsordning i slutet av 1990-talet. Närings- och handelsdepartementet (NHD) fick i uppgift att koordinera satsningarna och de olika åtgärderna. Genom eNorge-planen fick Kultur- och kyrkodepartementet (KKD) uppdraget att ta fram en handlingsplan för det norska språket och informations- och kommunikationsteknologin. Planen skulle ge riktlinjer för en språkpolitisk satsning på IKT-området och föreslå åtgärder för att realisera de uppsatta målen. Planen överlämnades till kulturministern i februari 2001.

En av de föreslagna åtgärderna var att skapa en norsk språkbank som kan användas i utvecklingen av språkteknologiska produkter för norska. Riktlinjerna för projektet fastställdes i en plan som lades fram för departementen (KKD och NHD) i oktober 2002. Projektet är ännu inte realiserat, men Språkrådet arbetar för närvarande (mars 2006) med att göra det möjligt att ta tillvara språkresurser från ett norskt språkteknologiskt företag som gick i konkurs 2003. Dessa resurser kan bli grunden för en norsk språkbank.

Både IKT-planen och arbetet med den norska språkbanken är resultatet av ett samarbete mellan Språkrådet, forskningsmiljöer och företag med intresse för språkteknologi. Det har gett Språkrådet ett mycket användbart och nödvändigt nätverk för vidare arbete.

Flera riksdagspropositioner har de senaste åren konstaterat att språkteknologin är en väsentlig del i det norska språkets möjlighet att överleva i framtiden, senast i kulturpropositionen som kom våren 2003. Ännu saknas dock konkreta politiska beslut som förbinder till att det satsas på detta område.

Forskningscentralen för de inhemska språken, Finland

Forskningscentralen har i en rad år samarbetat med företaget Lingsoft kring utvecklingen av programmet Svefix, ett stavnings- och grammatikkontrollverktyg baserat på motsvarande verktyg för MS Word. Programmet innehåller en finlandismkontroll för feltyper som särskilt

² www.språkteknologi.se

förekommer i svenska texter skrivna i Finland. Till planerna hör att ta fram versioner som fungerar i andra miljöer än MS Word för Windows.

Sedan mitten av 1990-talet har Forskningscentralen byggt upp en omfattande textbank med finska och svenska korpusar. Materialet omfattar över 200 miljoner ord finsk text och nästan 35 miljoner ord finlandssvensk text. I det finska materialet ingår också äldre språk och dialekter. De moderna korpusarna är tillgängliga för forskare via en extern server och kräver användartillstånd. Under år 2006 är det meningen att Forskningscentralen skall göra det diakrona materialet bättre tillgängligt genom att sätta upp en särskild korpusserver med webbgränssnitt för allmänheten. Arbete pågår med att enhetliggöra materialet.

Under 2005 har Forskningscentralen i samarbete med ett språkteknologiföretag tagit fram en finsk-svensk parallellkorpus på ca 4 miljoner ord, framför allt finländsk lagtext och diverse ekonomiska texter. Avsikten är att göra korpusen tillgänglig och sökbar för allmänheten. Korpusen har redan utnyttjats vid utvecklingen av en ny sorts översättningsprogram baserat på så kallad översättningsintelligens.

Den nyaste stora finska ordboken *Kielitoimiston sanakirja* utkom 2005. Ordboken finns endast i elektroniskt format och är resultatet av samarbete mellan Forskningscentralen och företaget Kielikone. Också Stora finsk-svenska ordboken som ges ut av förlaget WSOY i samarbete med Forskningscentralen föreligger i elektroniskt format.

Danska språknämnden, Danmark

Danska språknämnden är i färd med att göra en webbportal med språkresurser för det danska språket i samarbete med Danska språk- och litteratursällskapet. Språknämndens egen webbsida ska omarbetas i samband med detta så att båda webbplatserna använder samma CMS-system (Content Management System). Webbportalen beräknas vara klar i slutet av 2006.

Danska språknämnden planerar att införa ett nytt, XML-baserat redigeringsystem för Retskrivningsordbogen. Systemet ska också användas vid ordinsamlingen och svarsdatabasen. Svarsdatabasen ska innehålla språknämndens registrerade svar på språkfrågor och finnas tillgänglig via webbportalen. Retskrivningsordbogen på språknämndens webbplats har fått ny design och bättre presentation av sökresultaten.

Språksekretariatet, Grönland

Språksekretariatet har nyligen utvecklat webbaserade verktyg för att analysera grönländska ordformer och för att generera ordformer och hela böjningsparadigm³. Sekretariatet arbetar också med att ta fram skrivverktyg för MS Office-paketet vilka väntas vara klara i en betaversion i augusti 2006. Den grönländska nätordboken Katersat är under uppdatering och beräknas finnas i en ny version i slutet av 2006.

Språkteknologiska projekt i Nordens språkråds regi

Nordens språkråd har intresserat sig för språkteknologi och språkteknologins möjligheter i flera år. Det började med en kursändring i projektet Nordisk nätordbok som i juni 2003 i sin projektplan lade fram idén att ordböcker och andra språkliga resurser ska göras tillgängliga via en webbportal.

År 2004 utlyste Nordiska ministerrådet projektmedel för utveckling av en nordisk, flerspråkig sökmotor. Syftet var att göra Ministerrådets webbsidor sökbara på alla de skandinaviska huvudspråken oavsett vilket av språken som används i sökfrågan. Ett nordiskt konsortium skapades för uppdraget och resultatet väntas vara klart i juni 2006.

³ Se <http://www.oqaasileriffik.gl/>

Under arbetet med den nordiska nätordboken visade det sig nödvändigt att få klarhet i olika upphovsrättsliga bestämmelser i de nordiska länderna. En utredning påbörjades 2004. Nordens språkråd arbetar för närvarande på en projektplan och en kostnadsöversikt för en nordisk webbportal för språkområdet. Planen ska vara klar i april 2006.

Seminarium i Pargas

Tanken på en arbetsgrupp om språkvård och språkteknologi i Norden väcktes våren 2005 i Pargas, Finland på ett nordiskt seminarium om språkkontrollverktyg som finansierades av Nordens språkråd. Språknämnderna i Norden, språkteknologiska företag och organisationer samt enskilda forskare inom området inbjöds att delta i seminariet. En rapport från seminariet finns på arbetsgruppens webbplats⁴.

Seminariet kunde definiera ett antal uppgifter och problem som måste lösas om språkteknologin ska kunna tas i språkvårdens tjänst. De kan sammanfattas i åtta punkter:

- Vi bör definiera vad som utgör så kallade språkteknologiska resurser. Färdiga produkter bör inte betraktas som resurser, däremot korpusar, ordboksdata, öppen källkod, prototyper och grundläggande komponenter för lingvistisk analys, t.ex. morfologisk analys och disambiguering.
- På universitet och högskolor finns många befintliga resurser. Företag som gått i konkurs har också lämnat efter sig resurser och färdiga program. Tyvärr finns ingen utförlig sammanställning över dessa. Språknämnderna borde ta på sig uppgiften att sammanställa de resurser som redan finns.
- Språknämnderna och Nordens språkråd bör verka för att göra befintliga resurser tillgängliga för alla i lämpligt format. Forskare, statliga aktörer, språknämnder och utvecklare måste samarbeta för att få ut de resurser som redan finns.
- För de stora språken kan nya resurser tas fram av industrin eftersom det finns starka kommersiella intressen, men för relativt små språk som de nordiska behövs samhälliga medel. Därför bör språknämnderna och Nordens språkråd sätta press på regeringarna i syfte att ta fram nödvändiga resurser. När samhället delar ut ekonomiska medel för att utveckla nya resurser ska man ställa hårdare krav på att dessa ska bli allmänt tillgängliga. Ett juridiskt avtal som styr detta bör tas fram.
- Att resurser finns i många olika format är ett problem. Därför finns ett behov av en språkoberoende plattform med allmänt tillgängliga resurser i gemensamt format. Ju allmännare och mindre språkberoende, desto bättre. De nordiska länderna bör gå ihop om detta. Man bör satsa på öppen källkod så att resurser lätt kan vidareutvecklas och anpassas för nya språk och tillämpningar. Detta är inte minst viktigt för små språk som inte har stora ekonomiska medel för utveckling av språkteknologiska verktyg.
- Olika grupper har olika behov och möjligheter till finansiering. Vissa nordiska språk har få brukare och måste ges särskilt ekonomiskt stöd av samhället eftersom det kommersiella intresset för att utveckla språkteknologiska verktyg inte är tillräckligt stort. Vad gäller språkteknologi för skolorna finns ett visst kommersiellt intresse men språknämnderna har i det sammanhanget också en viktig uppgift att se till den språkliga kvaliteten på produkterna. Språknämnderna bör därför skaffa sig överblick över vilka produkter som finns till buds så att man om möjligt kan påverka producenterna att ta fram produkter anpassade för alla grupper, också grupper med speciella behov. Som myndighet kan språknämnderna inte obehindrat delta i kommersiell produktutveckling, men det behöver

⁴ www.nordisk-sprakrad.no/Sprakteknologisamarbete.htm

inte hindra att man samarbetar med utvecklarna vad gäller kraven på språklig kvalitet hos produkterna. Inom vissa ramar bör det också vara möjligt att bidra till att kraven realiseras t.ex. genom att bidra med ordböcker och språkregler.

- Det finns många svårigheter i samarbetet mellan forskning och programvaruindustrin. Juridiska problem med rättigheter, avtal och ersättning ställer till problem. Genom att samordna aktiviteter och etablera språkbanker med samhälleligt ägda resurser kan man underlätta problemen.
- Språknämnderna borde ta större ansvar, utarbeta krav och utvärderingsprinciper som definierar vad ett program ska klara och hur det ska utvärderas i språkligt hänseende. Kraven och principerna bör definieras tydligt, dokumenteras och göras tillgängliga för alla. Genom att följa dessa krav och principer får utvecklingsmiljöerna hjälp med att kvalitetssäkra sina produkter.

Seminariet kan förhoppningsvis bli startpunkten för ett nordiskt samarbete kring de uppgifter och problem som formulerats ovan. Mycket arbete återstår dock innan planerna kan konkretiseras och sättas i verket. De första stegen har nyligen tagits i och med bildandet av en språkteknologisk vismansgrupp och en arbetsgrupp för språkvård och språkteknologi i Norden.

En arbetsgrupp för språkvård och språkteknologi i Norden

På nätverksmötet för de officiella språkvårdsorganisationerna i Norden den 8 september 2005 i norska Bergen tillsattes en arbetsgrupp – Arbetsgruppen för språkvård och språkteknologi i Norden – med uppdrag att initiera ett samarbete om språkteknologi mellan språknämnderna i de olika länderna. Det långsiktiga syftet är att stärka språkvårdens inflytande över språkteknologiska produkter och deras konsekvenser för skriftspråksnormer och språkanvändning. Men arbetsgruppens uppgift är inte enbart övervakande. Lika viktigt är att informera om produkter som uppfyller god språklig kvalitet och att verka för att viktig teknik görs tillgänglig för språken i Norden, samt anpassas till grupper med särskilda behov.

Arbetsgruppen ska formulera en gemensam strategi för de nordiska språkvårdsnämndernas språkteknologiska arbete. Arbetets fokus ligger på frågor som rör normering av skrivet språk i språkteknologiska sammanhang. Det arbete som rör skrivverktyg som språkkontroll och elektroniska ordböcker kommer att prioriteras, som till exempel principer för språknormering i språkkontrollprogram. Ett grundligt utarbetande av sådana arbetsuppgifter väcker många svåra frågor, t.ex. vilka språkliga krav man kan ställa på språkteknologiska produkter av olika slag och på vilket sätt kvaliteten kan bedömas på ett rättvisande sätt. Till hösten 2006 planerar arbetsgruppen ytterligare ett seminarium för fortsatta diskussioner om det språkteknologiska samarbetet i Norden.

Till följd av seminariet i Pargas beslutade Nordens språkråd att tillsätta en vismansgrupp med uppdrag att ta fram ett förslag på språkteknologisk strategi för de nordiska länderna, och undersöka möjligheterna att etablera en gemensam nordisk språkbank. Arbetsgruppen för språkvård och språkteknologi i Norden bidrar med synpunkter på arbetet. Resultatet kommer att presenteras i den s.k. vismansrapporten som ska vara klar i juni 2006.

Summary

Human language technology products available in an individual's own language are crucial to achieving acceptance and deployment of the technology. The availability of products and services in the specific national language greatly facilitates the individual's ability to use the product. The language councils of the Nordic countries have focused on this matter in different ways, and have implemented different types of activities with regard to language quality in the technology. This article provides an overview of what the language councils

have done, what they are doing and which initiatives have been taken to enhance collaboration and coordination of their efforts.

Referenser

Bästa språket – en samlad svensk språkpolitik (2005). Proposition 2005/06:2. Utbildnings- och kulturdepartementet.

Domeij, R. (2005). *Datorn granskar språket*. Utg. av svenska språknämnden. Norstedts Akademiska förlag.

Handlingsplan for norsk språk og IKT. Norsk språkråd. Oslo 2001

Karlsson, O. (2004). *Skrivstöd för andraspråksskribenter – användaranpassning av språkkontrollen Granska för andraspråkssvenskar som skriver i arbetet*. Magisteruppsats i nordiska språk. Institutionen för nordiska språk, Stockholms universitet.

Karlsson, O. (2005). *Ändringsförslag till språkkontrollen i Microsoft Word*. Intern rapport. Svenska språknämnden.

Samling og tilgjengeleggjering av norske språkteknologiresursar. Norsk språkråd, Oslo 2002. (Prosjektplan for norsk språkbank)

Stortingsmelding nr. 48 (2002-2003): *Kulturpolitikk fram mot 2014*, Kultur- og kyrkjedepartementet, Oslo 2003.