

**HANDLINGSPLAN**

**FOR**

**NORSK SPRÅK OG IKT**

**Revidert utgave**

Norsk språkråd  
juni 2001

## Forord

Kulturdepartementet bad i brev av 14.07.2000 Norsk språkråd om å utarbeide en handlingsplan for norsk språk og informasjons- og kommunikasjonsteknologi (IKT). Planen er en oppfølging av St.meld. nr. 13 (1997-98) *Målbruk i offentlig teneste* og skal inngå i Nærings- og handelsdepartementets eNorge-plan. Språkrådet la 7. februar 2001 fram et forslag til handlingsplan. I forbindelse med at Kulturdepartementet er i ferd med å utarbeide en ny målbruksmelding, har det bedt om at tiltakene i forslaget blir gjennomgått på nytt, konkretisert og behørig forankret.

En referansegruppe med følgende sammensetning har bidratt med forslag, opplysninger og kommentarer under gjennomgåelsen:

AAD: rådgiver Siw Anita Vik  
KD: underdirektør Gro Paudal, førstekonsulent Eivind Lorentzen  
KUF: rådgiver Ellen Høy-Petersen  
NFR: avdelingssjef Bernt-Erik Heid, rådgiver Tron Espeli  
NHD: førstekonsulent Espen Arneberg Børset  
SHD: seniorrådgiver Inger Elisabeth Kvaase, rådgiver Inge Ovesen  
SIVA: assisterende direktør Tor Bjøru  
SND: kredittdirektør Bjørn Normann Hansen, rådgiver Kjell Haganæs

Arbeidet i referansegruppen har vært ledet av Språkrådet ved rådgiver Torbjørg Breivik. Også andre instanser og enkeltpersoner har kommet med synspunkter.

Språkrådet har behandlet den reviderte utgaven og legger med dette *Handlingsplan for norsk språk og IKT – Revidert utgave* fram for Kulturdepartementet.

Oslo, 20. juni 2001

Ola Haugen  
leder

Ståle Løland  
kontorsjef

# INNHold

|   |       |
|---|-------|
| <b>Kapittel 1 Sammendrag</b> .....                                      | s. 1  |
| <b>Kapittel 2 Bakgrunn</b> .....  | s. 2  |
| Kapittel 2.1 Innledning .....   | s. 2  |
| Kapittel 2.2 Bokmål – nynorsk .....                                     | s. 3  |
| Kapittel 2.3 Språk og teknologi .....                                   | s. 4  |
| Kapittel 2.4 Kultur, utdanning og forskning .....                       | s. 5  |
| Kapittel 2.5 Mål for norsk IKT-politikk .....                           | s. 6  |
| Kapittel 2.6 Nye utfordringer .....                                     | s. 6  |
| <b>Kapittel 3 Norsk språkteknologi</b> .....                            | s. 8  |
| Kapittel 3.1 Språkteknologi .....                                       | s. 8  |
| Kapittel 3.2 En norsk språkbank .....                                   | s. 9  |
| 3.2.1 Innholdet i språkbanken .....                                     | s. 10 |
| 3.2.2 Hva vil en norsk språkbank koste? .....                           | s. 10 |
| Kapittel 3.3 Utdanning i språkteknologi .....                           | s. 12 |
| Kapittel 3.4 Forskningsprogram for språkteknologi.....                  | s. 12 |
| Kapittel 3.5 Virkemiddel – næringsvirksomhet .....                      | s. 13 |
| <b>Kapittel 4 Norskspråklig programvare</b> .....                       | s. 15 |
| Kapittel 4.1 Norsk IKT-terminologi .....                                | s. 16 |
| Kapittel 4.2 Språket i programvare .....                                | s. 17 |
| Kapittel 4.3 Kontorstøtteprogram og administrativ programvare .....     | s. 18 |
| Kapittel 4.4 Elektroniske læremiddel .....                              | s. 18 |
| Kapittel 4.5 Teknisk programvare .....                                  | s. 20 |
| Kapittel 4.6 Nettsteder og underholdningsvarer .....                    | s. 21 |
| Kapittel 4.7 Norsk e-handel .....                                       | s. 22 |
| <b>Kapittel 5 Samarbeid og nettverksbygging</b> .....                   | s. 23 |
| <b>Kapittel 6 Konklusjoner</b> .....                                    | s. 25 |
| Kapittel 6.1 Handlingsplanen .....                                      | s. 25 |
| Kapittel 6.2 Tiltak etter kapittel .....                                | s. 25 |
| Forkortelser brukt i handlingsplanen .....                              | s. 28 |
| Aktuelle dokument/underlag for handlingsplanen .....                    | s. 29 |
| <i>Vedlegg 1: Treng vi ein norsk språkbank? Notat</i> .....             | s. 30 |
| <i>Vedlegg 2: Kostnadsoverslag for ein norsk språkbank. Notat</i> ..... | s. 33 |



# Kapittel 1 SAMMENDRAG

## Mål

1. I Norge skal IKT-produkt og -tjenester ha norskspråklig brukergrensesnitt som bygger på kunnskap om norsk tale og de to målformene
2. Alle med norsk som morsmål skal stilles likt i tilgang til og valg av norskspråklige IKT-produkt og -tjenester og skal gis like muligheter til å skaffe seg kompetanse i bruken av dem
3. Mengden norskspråklig programvare må økes, særlig programvare på nynorsk
4. Opplæringslovens bestemmelser om språklige parallellutgaver må oppfylles i praksis for elektroniske læremiddel
5. Språkteknologisk forskning og utvikling i Norge skal skje på norskspråklig grunn (bokmål, nynorsk og dialektene) og må tilføres tilstrekkelige økonomiske midler
6. En standardisert norsk IKT-terminologi må utvikles og stilles til rådighet
7. Norge skal delta aktivt i det nordiske og øvrige europeiske samarbeidet innenfor språkteknologi
8. Norge skal bidra til å styrke erfaringsutvekslingen mellom språkorganisasjonene i Norden og ellers i Europa om informasjons- og kommunikasjonsteknologiens innvirkning på nasjonalspråkene og om tiltak for å holde språkene i hevd også på dette området

## Tiltak

(*Ansvarlig* er den instansen Språkrådet anser som best egnet til å iverksette tiltaket.)

1. Etablere en norsk språkbank – ansvarlig: KD
2. Innføre et regelverk, med forankring i lov eller forskrift, for offentlige innkjøp av vanlig programvare (operativsystem medregnet) slik at det sikres språklige parallellutgaver – ansvarlig: AAD/KD
3. Gi økonomisk støtte til utvikling og bruk av verktøy for oversetting av programvare, også mellom bokmål og nynorsk – ansvarlig: KD
4. Gi økonomisk støtte til utvikling av elektroniske læremiddel på norsk (bokmål og nynorsk) og til bruk av norske språkteknologiprodukt i brukergrensesnittet deres – ansvarlig: KUF
5. Sørge for et fullverdig utdanningstilbud i språkteknologi – ansvarlig: KUF
6. Opprette et norsk forskningsprogram i språkteknologi – ansvarlig: NFR
7. Utrede behovet for en sektorovergrepene enhet som kan koordinere nasjonale tiltak innenfor språk og IKT – ansvarlig: Språkrådet
8. Samle, standardisere, utvikle og gjøre tilgjengelig norsk IKT-terminologi – ansvarlig: Språkrådet/RTT
9. Initiere, støtte og delta aktivt i tiltak knyttet til språk og IKT (på nasjonalt, nordisk og øvrig europeisk plan) – ansvarlig: Språkrådet

## Kapittel 2 BAKGRUNN

### 2.1 INNLEDNING

I St.meld. nr. 13 (1997-98) *Målbruk i offentlig teneste* rettes søkelyset mot norsk språk og informasjons- og kommunikasjonsteknologi (IKT). Kultursektoren har et spesielt ansvar for å stimulere til norskspråklig innholdsproduksjon på IKT-området. I kapittel 4 i stortingsmeldingen gir Kulturdepartementet i følgende målformulering uttrykk for regjeringens bekymring for norsk språks stilling i det framvoksende informasjonssamfunnet: *”Det trengs eit offensivt handlingsprogram der både juridiske, administrative, språkteknologiske og økonomiske verkemiddel blir tekne i bruk. Målet er å kunne gi både nynorsk- og bokmålsbrukarane gode reiskapar til støtte for språkbehandling og å tilby relevante informasjonstenester o a på begge målformer”* (s. 29). I Innst. S. nr. 174 (1997-98) er bekymringen enda tydeligere: *”Det vil lett kunna bli slik at dersom ikkje staten tek den kulturpolitiske styringa når det gjeld språk og IT, vil tilfeldige programvareutviklarar kunna leggja føringer på utviklinga av norsk språk.”* (s. 1 i *Merknader*)

Regjeringen har i sin eNorge-plan (v. 1.0 og 2.0) satt seg som mål å skape et informasjons- og kunnskapssamfunn for alle. eNorge er regjeringens handlingsplan for bruk av elektroniske verktøy og hjelpemiddel på ulike områder i samfunnet samt innhold i elektroniske tilbud og tjenester. Planen er en parallell til EUs handlingsplan eEurope. Språkrådets handlingsplan for norsk språk og IKT inngår i eNorge. Språkrådet mener at tiltakene i handlingsplanen vil bidra til å virkeliggjøre regjeringens mål. Skal alle få tilgang til, kompetanse i og tillit til den nye teknologien og de nye mediene, må det norske språket brukes aktivt på alle områder innenfor IKT.

Stadig flere av oss tar i bruk elektroniske innretninger som kommunikasjonsmedium. I internasjonalt samarbeid og internasjonal kommunikasjon og handel er engelsk det dominerende språket. Informasjons- og kommunikasjonsteknologi er den fremste spredningsveien for engelsk som bruksspråk. Dette må vi ta konsekvensen av ved å iverksette styrkingstiltak som gjør brukerne oppmerksomme på faren for språklig domenetap og verdien av å opprettholde norsk (bokmål og nynorsk) som et fullgodt bruksspråk på alle områder.

Mulighet for deltakelse i offentlig debatt og medbestemmelse i beslutninger som treffes, er en forutsetning for et demokrati som det norske. Den teknologiske utviklingen innebærer nye utfordringer og gir samtidig nye muligheter. Godt og lett tilgjengelig språk fremmer god kommunikasjon, noe som er en forutsetning for innsyn og deltakelse. IKT-tjenester som er tilrettelagt slik at bruken blir mer intuitiv, vil lette tilgangen til informasjon for en stor del av befolkningen og øke muligheten for at flere kan delta i samfunnsutviklingen. I sitt program for fornyelse av offentlig sektor tar regjeringen sikte på å gjøre forvaltningen mer tilgjengelig gjennom økt bruk av IKT. Målet er at tjenestene som tilrettelegges ved hjelp av ny teknologi, skal være tilgjengelige for hele befolkningen.

Internasjonalt har fag- og forskningsmiljøer benyttet datamaskiner til utforskning av språk i om lag 40 år. Det var imidlertid først på 1990-tallet at man ble opptatt av den innvirkning IKT kan ha på språkbrukerne og språkbruken deres. EU satte i gang en rekke forsknings- og utviklingsprogram i siste halvdel av tiåret for å styrke kompetanse- og produktutviklingen i medlemslandene og for å møte utfordringene fra USA og Asia. I mange land har man vært opptatt av det økende presset som små språksamfunn er utsatt for. Det er bred enighet om at nasjonalspråket/-ene er grunnsteinen i et lands kultur. Språkteknologiske produkt på eget

språk blir derfor et betydningsfullt satsingsområde. Et viktig resultat av prosjektene under EU-programmene er opprettelsen av organisasjoner som har spesialisert seg på utvikling av standarder for bearbeiding, lagring og formidling av språksamlinger og termbaser. For en norsk språkbank vil det bli naturlig å knytte kontakt med slike organisasjoner. I flere europeiske land er det for øvrig i løpet av det siste tiåret etablert språkbanker (England, Finland, Frankrike, Tyskland). I andre land pågår det innsamling av materiale med sikte på opprettelse av nasjonale språkbanker (f.eks. flamsk tale i Belgia og Nederland).

EU-programmene bygger på en prosjektmodell som krever at det er deltakere fra minst tre land og komplementær kompetanse og faglig bakgrunn representert i prosjektgruppen. Hensikten har vært å kople forskningsmiljø, næringsliv og sluttbrukermiljø.

I denne handlingsplanen pekes det på områder der Norge kan delta i utviklingen og utnytte synergieffektene som oppstår ved tverrfaglig arbeid og samordnet innsats. Planen tar utgangspunkt i at norsk fortsatt skal være vårt felles kultur- og arbeidsspråk, og i erkjennelsen av at norsk er et fullgodt språk som kan brukes på alle områder og i alle fag. For å virkeliggjøre ambisjonene om å gjøre IKT til et konkurransefortrinn for norsk næringsliv kreves det tverrfaglig tenkning og nye fagkombinasjoner. Skal flest mulig få tilgang til og kunne nyttiggjøre seg den nye teknologien, er norskspråklig programvare en forutsetning. En allmenngjøring av informasjons- og kommunikasjonsteknologien betyr også at flere kan bidra i teknologi- og produktutviklingen, noe som i neste omgang medfører kompetanseheving og -spredning. Handlingsplanen peker på hva som må til på ulike områder, hvilke forutsetninger som må oppfylles, hvilke sammenhenger det er mellom enkelte av områdene, og hvem Språkrådet mener bør ha ansvaret for å sette i verk de ulike tiltakene.

## **2.2 BOKMÅL – NYNORSK**

Norsk har to offisielle, likestilte målformer, og det er bred politisk enighet om at det fortsatt skal være slik. På IKT-området er situasjonen svært ulik for bokmål og nynorsk. De fleste dataprogram som tilbys på norsk, foreligger gjerne bare på bokmål. Nesten ingen program er opprinnelig utviklet med nynorsk grensesnitt. Noen få foreligger på begge målformer, men det er langt igjen før enkeltpersoner, skoler, kommuner, fylkeskommuner, statsorgan og virksomheter som har nynorsk som førstevalg, får alle programmene de har behov for, på sin målform. Den språklige kvaliteten i dataprogram som er oversatt til norsk, varierer fra ganske bra til så dårlig at originalspråket skinner tydelig gjennom. Dette gjelder også for program oversatt fra bokmål til nynorsk. Iblant har man ikke engang tatt seg bryderiet med å oversette all tekst i programmets brukergrensesnitt.

Regjeringen slo med St.meld. nr. 13 (1997-98) fast at den ønsker å endre denne situasjonen. Et flertall på Stortinget (Ap, KrF, Sp og SV) fulgte opp regjeringens holdning, bl.a. med følgende uttalelse: *"Dette fleirtallet har merka seg at Regjeringa går inn for eit offensivt handlingsprogram der både juridiske, administrative, språkteknologiske og økonomiske verkemiddel vert tekne i bruk. Målet må vere å kunna gi både nynorsk- og bokmålsbrukarane gode reiskap til støtte for språkbehandling og tilby tenlege informasjonstenester o.a. på begge målformer."* (Innst. S. nr. 174 (1997-98), s. 1 i *Merknader*)

Det må tas i bruk flere typer virkemiddel for å nå dette politiske målet. Som en av landets største IT-kunder rår det offentlige samlet over et sterkt økonomisk virkemiddel til å kreve at produsentene utvikler nynorskversjoner. Norsk språkråd mener for det første at det er

nødvendig å endre regelverket for offentlige innkjøp slik at det blir et krav til leverandørene at de må tilby produktene på både bokmål og nynorsk. En økt bevissthet blant nynorskbrukere om betydningen av å opptre samlet som en stor aktør i markedet vil være et godt pressmiddel overfor leverandørene. I tillegg kan det offentlige gi økonomisk støtte til å videreutvikle oversettingsverktøy som kan forenkle arbeidet med å lage programversjoner på begge målformer. Samarbeid med bransjen om tiltak og prosjekt som kan føre til bedre språk generelt (terminologiutvikling og andre språkfaglige tjenester), vil også bidra til å gjøre det lettere for produsentene å imøtekomme kravet om språklige parallellutgaver (se også kap. 4.4).

Språkrådet gjorde i januar 1999 og 2000 systematiske søk i informasjonstilfanget på informasjonstjenestene Alta Vista og ODIN på Internett. Resultatet viste 5-6 % nynorsk på Alta Vista og 10-11 % nynorsk på ODIN. ODIN er departementenes informasjonssystem og skal oppfylle kravene til målbruk i offentlig tjeneste (minst 25 % på hver målform). En mangel ved tjenestene på ODIN er at man må søke enten på bokmål eller på nynorsk. Hadde tjenestene ved søking gitt tilslag for samme ord i begge målformer, ville det økt nytteverdien for brukerne.

Norsk skriftspråk, både bokmål og nynorsk, har mange valgfrie former som det må tas hensyn til i norskspråklig programvare som gjør bruk av språkteknologi. Mange oppfatter valgfriheten som en berikelse av språket og et uttrykk for språklig og sosial toleranse. Andre betrakter valgfriheten som en ulempe og mener at en fastere skriftnorm ville gagne alle språkbrukere. Valgfriheten er viktig og nødvendig i dagens språkpolitiske situasjon og kan ikke innskrenkes gjennom bruk av språkteknologiske hjelpemidler uten at det får uheldige språkpolitiske konsekvenser. Det er ikke ønskelig at teknologien på denne måten skal styre språkpolitikken. Teknisk sett kompliserer og fordyrer valgfriheten i noen grad utvikling og produksjon av norskspråklig programvare som gjør bruk av språkteknologi. Samtidig ligger det interessante faglige utfordringer i å framstille avanserte programtekniske løsninger som ivaretar valgfriheten. Slike løsninger må ikke begrense språkbrukernes valgmuligheter. Språkrådet vil overvåke utviklingen på dette området og om nødvendig ta initiativ for å lede den i ønsket retning. Språkrådet kan ikke godta at teknologien virker begrensende på språkbruken eller bestemmende for språkutviklingen, eller at industrien utformer løsninger som innsnevrer brukernes valgfrihet. Stavekontrollprogram kan gi brukeren mulighet til å gjøre personlige formvalg. Valgene kan gjøres grove og enkle for dem som ikke er spesielt opptatt av formvalget, og detaljerte for dem som legger vekt på det individuelle. Her er det altså snakk om funksjoner som brukeren kan slå av og på etter eget ønske. Et godt stavekontrollprogram kan i tillegg til sitt hovedformål også virke som en ”rettskrivningsveileder” som gir innsikt i norsk språk. Språkrådet vil bidra til utvikling av skrivestøtteprogram med slik funksjonalitet.

## 2.3 SPRÅK OG TEKNOLOGI

Det er et nasjonalt ansvar å legge best mulig til rette for utvikling av norskspråklig programvare og verktøy på norsk på alle IKT-områder. Det dreier seg om å utvikle norske program, lage norskspråklige utgaver av utenlandske program, videreutvikle verktøy for oversetting og språkversjonering<sup>1</sup>, utvikle norsk terminologi, utvikle norskbaserte verktøy for

---

<sup>1</sup> Språkversjonering: Det å lage forskjellige utgaver av et dataprogram for ulike varianter av ett og samme språk eller for ulike språk ved å oversette alt språk (tekst og/eller tale) i brukergrensesnittet, også meldinger, brukerveiledning og hjelpetekster.



skrivestøtte (tilleggsprogram for orddeling, stavekontroll, grammatikkontroll, stilanalyse, tekstanalyse og tekstsammendrag, dikteringssystem osv.) og norskbaserte taleteknologiske produkt. Disse hjelpemidlene kan bare realiseres i krysningpunktet mellom teknologisk kompetanse og kunnskaper om norsk språk. I EU legges det stor vekt på utvikling innenfor IKT og språkteknologi. EUs MLIS-program ble opprettet nettopp for å rette søkelyset mot slik utvikling. Målet med programmet var å styrke bruken av ulike teknologier, verktøy og metoder bl.a. ved å støtte opprettelse av sentre for språkressurser og effektiv bruk av disse. Et annet EU-program, INFO2000, skulle bidra til innsamling og tilrettelegging for bruk og gjenbruk av materiale egnet som multimedie-innhold. Programmet eContent viderefører målene i både MLIS og INFO2000, som begge er utløpt. I eContent er oppmerksomheten i større grad rettet mot innholdet som sådant, der det skjer en kopling mellom språk og kultur, enn mot tjenestene eller produktene som formidler innholdet.

Språkteknologiske produkt er dels språkuavhengige, dels språkbundne. Det betyr at de bygger både på generell teknologi og på teknologi som forutsetter særpråklige spesialrutiner på grunnlag av systematisk og strukturert analyse av det aktuelle språket. Skal man utvikle grammatikkontroll for norsk, sier det seg selv at man ikke kan bruke regler for noe annet språk. Regler for dansk og svensk kan til en viss grad tilpasses norsk og omvendt, men like fullt må utviklingen bygge på kunnskaper om det enkelte språk. Det er opplagt for alle at taletjenester i Norge må ta utgangspunkt i norsk språk, men ikke alle er klar over at det må store mengder talemateriale til for å utvikle f.eks. dikteringssystem.

Språkteknologien blir et viktig arbeidsfelt i årene som kommer. Språkteknologisk forskning og produktutvikling må i Norge bygge på kunnskap om det norske språket. Skal vi oppfylle de politiske målene på IKT-området, må det utdannes flere med språkteknologisk kompetanse. Det må dessuten forskes videre i krysningpunktet mellom teknologi og språk, og det må finnes risikovillig kapital til nyetableringer. Alle tendenser peker i retning av at nye tjenester og nye kombinasjoner av tjenester integreres i samme medium. Brukerne får tilgang til verktøy som lar dem selv velge og sette sammen tjenester etter behov. Utfordringen for Norge er å satse tilstrekkelig til at det kan tilbys norske produkt og norskspråklige tjenester innen rimelig tid. Et nært samarbeid mellom næringslivet og utdannings- og forskningssektoren er påkrevd for å virkeliggjøre målet.

## **2.4 KULTUR, UTDANNING OG FORSKNING**

Innenfor IKT skjer utviklingen raskt. Tilegnelse av tekniske kunnskaper på området foregår gjerne på engelsk. Allmenngjøring av kunnskaper om og kompetanse i bruk av teknologien som verktøy og hjelpemiddel forhindres eller forsinkes til en viss grad når brukerprogrammene foreligger bare på engelsk. Norsk er et fullgodt språk som skal kunne brukes på alle områder og i alle fag. Det krever at man må arbeide systematisk og sette i verk konkrete tiltak for å hindre domenetap for det norske språket på IKT-området.

Norske bibliotek og mediebedrifter råder over store mengder materiale på norsk som kan inngå som innhold i norske bredbåndstjenester. Myndighetene og rettighetshaverne må sammen finne løsninger som sikrer tilgang til og gjenbruk av disse ressursene. Klarer man å komme fram til hensiktsmessige løsninger, er det et godt utgangspunkt for å sikre norskspråklig innhold og norsk kulturarv i de nye digitale nettverkene.

Kirke-, utdannings- og forskningsdepartementets satsing på IKT i utdanningssystemet forutsetter at det er tilgjengelig norskspråklig programvare og norskspråklige nettressurser. Målet er at alle elever og studenter uansett nivå og fagvalg skal bruke IKT-produkt og -tjenester som hjelpemiddel. Universitets- og høyskolesektoren har på sin side prioritert IKT-faglige utdanningsløp og økt antall studieplasser på området gjennom hele 1990-tallet, dels innenfor foreliggende rammer og dels som følge av at KUF har lagt føringer i den retning i forbindelse med utvidelse av rammene for antall studieplasser (jf. også *IKT i norsk utdanning. Plan for 2000-2003*, KUF 2000).

## 2.5 MÅL FOR NORSK IKT-POLITIKK

I planen *Norge – en utkant i forkant* (NHD, februar 1998) uttaler regjeringen at målet er å gjøre IT (IKT) til et av de største konkurransefortrinnene for norsk næringsliv. Dette forutsetter avansert bruk av informasjonsteknologi i norsk næringsliv, noe som i sin tur kan føre til nye arbeidsmåter og nye organisasjonsformer. Kravet om og nødvendigheten av økt internasjonalisering skaper større behov for tverrfaglig kompetanse. Språk- og kulturkunnskap i kombinasjon med annen kompetanse blir stadig viktigere (s. 9), og det vil være nødvendig å ha fagkunnskap i både språk og informasjonsteknologi.

I NHDs eNorge-plan, versjon 1.0, følges dette opp i avsnittet *Individ, kultur og miljø*, som har som hovedmål at ”vi skal legge grunnlaget for et informasjons- og kunnskapssamfunn for alle” (s. 4). Budskapet er at flest mulig skal settes i stand til å nyttiggjøre seg den nye teknologien, og at det er nødvendig med en innsats for å gjøre nye produkt mer brukervennlige. Det blir videre pekt på at det kreves en samordnet innsats for å sikre bruk av norsk språk og motvirke den dominerende stillingen engelsk har i IKT-næringen.

Norges forskningsråd framhever i sin handlingsplan for IKT-forskning at det må skje en omfattende heving av den generelle IKT-kompetansen i samfunnet dersom man skal nå målet om at Norge i 2010 skal være regnet blant verdens ledende IKT-nasjoner. NFR ser for seg en tredobling av bevilgningene til forskning i IKT i løpet av de påfølgende fem årene. I tillegg pekes det på at det er nødvendig med grunnleggende infrastrukturiltak som f.eks. etablering av en norsk språkbank.

## 2.6 NYE UTFORDRINGER

Tverrfaglig samarbeid vil være en avgjørende suksessfaktor for satsingen på IKT. I utdanningen av fagfolk må fag som tidligere ikke har samarbeidet i særlig grad, inngå et nærmere samarbeid for å utvikle nødvendig ny kunnskap. I industrien trengs det kvalifisert personell med tverrfaglig kompetanse. Språkvitere trengs på lik linje med teknologer for å framstille gode språkteknologiske produkt. I utviklingen av velegnede brukergrensesnitt er det behov for fagfolk med kunnskap om bl.a. psykologi, ergonomi, språk og teknologi. Næringslivet vil ha nytte av tettere samarbeid med forskningsmiljøene både for å få innsikt i hva som skjer på forskningsfronten, og for å kunne formidle hvilke behov det selv har. Det er likevel ikke slik at et forskningsresultat (en prototyp) uten videre kan gjøres om til et salgbart produkt, men erfaringen og kunnskapen som ligger bak resultatet, kan danne et nyttig og fruktbart utgangspunkt for samarbeid og mulig videre utvikling.

Et nordisk samarbeid innen språkteknologi kan være kostnadseffektivt og givende for både forskere og næringsutviklere. Mye av det grunnlagsarbeidet som må gjøres for norsk, vil være tilsvarende for dansk og svensk. Man vil kunne dra nytte av at strukturen i de skandinaviske språkene ikke er så ulik. Verktøy utviklet for et av språkene vil lettere kunne tilpasses de andre enn verktøy utviklet for f.eks. engelsk. Eksempel på at man ser store muligheter for nordisk samarbeid, er etableringen av Nordisk Språkteknologi AS (NST) og SAIL Port Northern Europe AS (SPNE) på Voss. NST ble opprinnelig finansiert av norske og belgiske aktører og har benyttet belgisk teknologi som utgangspunkt for produkt- og tjenesteutviklingen. Selskapet er i dag 98 % norskeid. NST har språkressurser for dansk, norsk og svensk, og produktene rettes inn mot markedet i de tre landene. SPNE, som nå er eid av SIVA, ble etablert med betydelige midler fra NHD og har som formål å være ”fødselshjelper” for nordisk næringsvirksomhet innenfor språkteknologiske produkt og tjenester. Det vil være mye å vinne på utveksling og tilpasning av materiale også i arbeidet med å utvikle og standardisere IKT-terminologi.

Internasjonalt har forskere alltid samarbeidet på sine respektive fagfelt, utvekslet erfaringer og drevet prosjekt sammen. Norske språkressurser som gjøres tilgjengelig for gjenbruk, vil utfylle den europeiske satsingen på språkteknologisk forskning og produktutvikling. Ved utvikling av oversettingsverktøy og elektroniske taletjenester for flere språk vil flerspråklige korpus være til stor nytte på tvers av landegrensene og utgjøre et nødvendig materiale for forskere og utviklere i industrien.

## Kapittel 3 NORSK SPRÅKTEKNOLOGI

### MÅL

Språkteknologisk forskning og utvikling i Norge skal skje på norskspråklig grunn (bokmål, nynorsk og dialektene) og må tilføres tilstrekkelige økonomiske midler

Språkteknologiske produkt og tjenester, språkteknologisk programvare og andre IKT-produkt til allmenn bruk i Norge skal ha et norskspråklig brukergrensesnitt

Norsk språkteknologisk forskning og utvikling skal være i fremste rekke i Norden og internasjonalt

### TILTAK – norsk språkteknologi (SP):

| Nr. | Tiltak   | Ansvarlig | Samarbeidsparter                   | Ref./komm.      |
|-----|--|-----------|------------------------------------|-----------------|
| SP1 | Etablere en norsk språkbank  | KD        | Språkrådet, NFR, NHD               | Utredet i 2001  |
| SP2 | Sørge for et fullverdig utdanningstilbud i språkteknologi  | KUF       | Universitets- og høyskolesektoren  |                 |
| SP3 | Opprette et forskningsprogram i språkteknologi   | NFR       | KUF, NHD, KD                       | Er under arbeid |
| SP4 | Fremme industriell utnyttning av norsk språkteknologi  | NHD       | NFR, SND, SIVA                     |                 |
| SP5 | Utrede behov for og legge til rette språkteknologiske produkt og tjenester beregnet på brukere med særlige behov | SHD       | KUF, NHD, NFR (IT-FUNK, TYIN), AAD |                 |

### 3.1 SPRÅKTEKNOLOGI

I stadig økende grad må eller kan skrift og tale brukes når mennesker skal nyttiggjøre seg elektroniske innretninger og maskiner. Språkteknologiske produkt er hjelpemiddel eller verktøy som forenkler gjøremål i arbeidsliv og på fritiden samt kommunikasjon mennesker imellom.

Språkteknologi er enkelt sagt all teknologi som bygger på systematisert kunnskap om og bruk av naturlig språk. Når f.eks. husholdningsmaskiner styres ved hjelp av tale, inneholder de et lite språkteknologisk program som gjenkjenner lydsignalene og ”oversetter” dem til en kommando, som så utføres. Tekstbehandlingsprogram inneholder vanligvis bl.a. en stavekontroll som bygger på en elektronisk ordliste. Ord brukeren skriver feil, markeres og kan rettes. Grunnord og bøyningsformer som finnes i ordlista, vil ikke bli markert så lenge de er riktig skrevet, men blir brukt feilaktig. Slike grammatiske feil vil bare avdekkes om

tekstbehandleren inneholder et "intelligent" språkteknologisk tilleggsprogram, en grammatikkontroll.

De søkesystemene som finnes i dag, er ikke gode nok for norsk språk. Ingen av dem tilbyr samtidig søk på bokmål og nynorsk på nettsteder og/eller i baser. Brukeren må ofte søke én gang for hver målform, og iblant flere ganger dersom en av eller begge målformene har valgfrie former. Det fins noen program for maskinstøttet oversetting, men kvaliteten og nytteverdien varierer mye. Utvikling av egnede oversettingsverktøy kan være til stor nytte i arbeidet med å øke andelen norskspråklig programvare og, ikke minst, bidra til utgivelse av språklige parallellutgaver av generell programvare.

Dikteringssystem som kan overføre tale til tekst, er et annet eksempel på språkteknologiske produkt som kan bli svært nyttige og kostnadssparende i visse sammenhenger. Det utvikles nå en norsk prototyp som kan benyttes til journalføring på sykehus. Kostnadsoverslag viser at om man oppnår 5 % innsparing på dette området, utgjør det tosfrede millionbeløp. Et tredje aktuelt eksempel er system som oversetter tekst til tale (syntetisk tale) og talestyrte hjelpemiddel for ulike formål. Språkteknologien vil gjøre informasjons- og kommunikasjonsteknologien mer tilgjengelig for folk flest og kan bli til stor hjelp for mange grupper som har særlige behov (små barn, eldre, funksjonshemmede, dyslektikere m.fl.). Blant annet har et samarbeidsprosjekt mellom NTNU og SINTEF i Trondheim resultert i et skrivestøtteprogram for dyslektikere (LingDys). Den største språkteknologiske satsingen i norsk industriell sammenheng er Nordisk Språkteknologi AS på Voss. Selskapet har utviklet system for talegjenkjenning og talesyntese for ulike formål på dansk, norsk (bare bokmål) og svensk. NST står for utviklingen av den språkteknologiske komponenten i den førnevnte prototypen for elektronisk journalføring i helsesektoren.

Betydningen av språkteknologien i morgendagens samfunn er blitt framhevet av mange. Microsofts Bill Gates har gjentatte ganger understreket den sentrale rollen taleteknologien vil ha framover i alle typer tekniske innretninger. Det er viktig at flest mulig blir i stand til å utnytte de muligheter IKT-utviklingen tilbyr. Språkteknologien bidrar til å senke brukerterskelen og gjør flere i stand til å benytte seg av IKT-baserte verktøy og informasjonstjenester.

Språkteknologien er av natur språkavhengig. Det er arbeids- og kostnadskrevenende å samle inn språkressurser og utvikle språkteknologiske produkt og tjenester for et språk. Ønsker vi norskspråklige produkt og tjenester av god kvalitet, trenger vi betydelige mengder norske språkressurser. Arbeidet med å framskaffe og tilrettelegge disse forutsetter både teknologisk og lingvistisk kompetanse.

## **3.2 EN NORSK SPRÅKBANK**

Språkteknologisk forskning og utvikling krever store mengder språkdata ("språkressurser"). Det samme gjelder for tilpasning av utenlandske språkteknologiprodukt til norsk språk. Vi har to offisielle målformer i Norge, og dialekt benyttes i alle muntlige sammenhenger. For å utvikle norskspråklige språkteknologiprodukt trengs det derfor minst like store norske språkressurser som for verdensspråket engelsk. Dersom vi ønsker at norsk språk skal holdes i hevd på alle områder i det framvoksende informasjonssamfunnet, er det nødvendig å bygge opp en stor samling norske språkressurser i en nasjonal språkbank. En slik språkbank vil være et verdifullt bidrag til arbeidet med å ta vare på variasjonsbredden i norsk skrift- og

talespråkskultur. Det vidtfaavnende språkmaterialet vil bl.a. være en forutsetning for utvikling av brukervennlig språkteknologisk programvare som ivaretar hensynet til den språklige valgfriheten her i landet.

### **3.2.1 Innholdet i språkbanken**

En språkbank må inneholde store mengder tekst (bokmål og nynorsk) og tale ((hoved)dialekter og normalisert uttale) som er samlet inn, systematisk bearbeidet og tilrettelagt for gjenbruk. Det trengs også ordlister, terminologisamlinger, emnekataloger, navnelister osv. All tekst og tale må merkes med informasjon om sjanger, opphav(spersion), brukssituasjon, opptaksdato o.l. Gode konversasjonsleksikon og/eller tesauruser vil også være være av stor verdi for en språkbank. Språkbanken må være en kontinuerlig virksomhet som løpende tilføres nytt materiale. Nytteverdien for språkteknologisk forskning og, ikke minst, produktutvikling er størst når språkdataene i banken er representative for samtidens språkbruk. Materiale som blir for gammelt for produktutviklingsformål, vil likevel fortsatt ha stor verdi for språkforskere.

En norsk språkbank – et nasjonalt korpus – ble utredet i 1999 (Svendsen 1999). Temaet ble diskutert i det norske forsknings- og utviklingsmiljøet. Det var bred enighet i miljøet om at det burde etableres en språkbank. Rapporten inneholder en oversikt over eksisterende norske språkdata og en liste over prioriterte suppleringer. For at industrien skal kunne utvikle robuste brukersystem på dagens internasjonale nivå, trengs det større mengder språkdata enn det som er nevnt i oversikten i rapporten, og det avgjørende momentet vil være bearbeiding og tilrettelegging av dataene. Rapporten drøfter også problemstillinger knyttet til forvaltning, organisering og finansiering av en språkbank. Det er naturlig å legge denne rapporten til grunn for det videre arbeidet. Ut fra et brukerperspektiv er det viktig å vurdere hvilke språkressurser som er nødvendige når norsk språk (bokmål og nynorsk) skal inngå i flerspråklige produkt som søkesystem og oversettingsverktøy. Det må dessuten sørges for at de behov språkteknologiindustrien har, imøtekommes. Likeledes må det tas hensyn til de krav som stilles til dagens språkteknologiske produkt og tjenester for andre språk.

Før man kan etablere en norsk språkbank, må følgende gjøres:

- a) Utrede juridiske, organisatoriske og datatekniske problemstillinger
- b) Oppnevne et interimsstyre for Norsk språkbank
- c) Lage kravspesifikasjon med oversikt over materiale som finnes og kan brukes, eller som mangler (interimsstyret er sentralt i denne sammenheng)
- d) Utarbeide finansieringsplan for Norsk språkbank

Kulturdepartementet er innforstått med den store betydning fagmiljøene tillegger språkbanken, og med behovet for utredninger i den forbindelse. Sentrale instanser som har uttalt seg om tiltaket, ser det som viktig å starte utredningene snarest mulig slik at etableringen av språkbanken ikke utsettes lenger enn nødvendig. Arbeidet med å oppnevne et interimsstyre, utarbeide kravspesifikasjon og finansieringsplan er bare til en viss grad avhengig av utredningene under pkt. a).

### **3.2.2 Hva vil en norsk språkbank koste?**

Utviklingskostnadene for språkteknologiske produkt og tjenester er tilnærmet like uansett språk, mens inntjeningspotensialet er avhengig av antall brukere. Det betyr at offentlig engasjement er nødvendig for å utvikle språkteknologiske produkt for og drive språkteknologisk forskning i et lite utbredt språk som norsk. Særlig vil det kreve et stort økonomisk løft å samle inn og bearbeide en tilstrekkelig mengde språkdata til å kunne drive

slikt forsknings- og utviklingsarbeid. Å samle språkdata i et nasjonalt korpus som gjøres tilgjengelig for alle aktører innenfor språkteknologisk FoU, er uansett en atskillig mer effektiv utnyttning av offentlige midler enn å gi tilskudd til samme formål i forbindelse med enkeltstående forskningsprosjekt og produktutvikling på området.

Den utbredte bruken av dialekter i Norge fører til at talematerialet må være større enn for mange andre språk. Den store valgfriheten i hver av de to skriftnormene vil også bidra til høyere kostnader for norsk. Utviklingsarbeidet blir av samme grunn mer omfattende. Språksituasjonen gjør at man i Norge får sammenfallende interesser mellom språkforskning og industriell FoU i kravet til variasjon i innholdet i språkbanken. Det finnes omfattende offentlige og private språkressurser som man kan begynne med. De må gjennomgås og vurderes i lys av kravene til gjenbruk, og man må sette eventuell bearbeiding av dem opp mot industriens og forskningens behov for nytt og aktuelt materiale. Det er viktig at man ikke bruker tid og midler på bearbeiding av materiale som ingen har umiddelbar nytte av, men tar utgangspunkt i de behov og krav brukerne av språkbanken og språkbrukerne har. Det er også viktig å ta i betraktning de høye kostnadene som den særnorske språksituasjonen medfører. Oppgaven med å etablere en språkbank er for stor for en enkelt institusjon eller virksomhet. Det må et nasjonalt løft til for å etablere en språklig fellesressurs som en språkbank.

En språkbank må ha et styre, en driftsorganisasjon, en administrasjon og et innhold. Arbeidet med å etablere en språkbank har tre hovedfaser:

- a) Utredning av juridiske, organisatoriske, tekniske og faglige forhold, ca. 0,5 mill. NOK
- b) Etablering av en driftsorganisasjon og et produksjonsapparat, ca. 6 mill. NOK
- c) Innsamling og tilrettelegging av språkmateriale til språkbanken, ca. 60-80 mill. NOK

Det tas forbehold om at det er mange usikkerhetsfaktorer i grunnlaget for oppstillingen av kostnadsoverslaget ovenfor (se vedlegg 1 og 2). At det dreier seg om en stor investering, er det imidlertid ingen tvil om. Innsamlingen og tilretteleggingen av språkmaterialet er tid- og arbeidskrevende. Oppbyggingen av språkbankens nødvendige minsteinnhold vil måtte foregå over et visst antall år. Det er vanskelig å fastslå nøyaktig hvor mange år arbeidet vil ta. Tilgang på kvalifisert arbeidskraft vil være den mest avgjørende faktoren. Når arbeidet med minsteinnholdet er avsluttet, må språkbanken løpende tilføres nytt, samtidig materiale. Anslaget i pkt. b) gjelder bare det første driftsåret. Påfølgende års utgifter til løpende drift, vedlikehold og videre utvikling er ikke vurdert i denne omgang fordi de er avhengige av bl.a. valg av driftsorganisasjon og finansieringsprinsipp for bankens tjenester.

Tilgang på språkdata vil kunne styrke framveksten av mangfoldig norsk språkteknologisk virksomhet i industriell regi, herunder bedrifter som framstiller generelle språkteknologiprodukt. Slike språkdata vil gi utviklere av IKT-baserte produkt og tjenester mulighet til å tilføre sine løsninger norskspråklig brukergrensesnitt. På grunn av de høye kostnadene forbundet med innsamling og bearbeiding av språkdata vil fri eller rimelig tilgang til slike data være en forutsetning for etablering av ny virksomhet og et variert tilbud av produkt og tjenester. Tilgang på språkdata er i tillegg en forutsetning for å kunne drive forskning på høyt nivå ved universitet og høyskoler. En høy forskningskompetanse og uteksaminering av tilstrekkelig mange kandidater med kompetanse i språkteknologi er nødvendig for å kunne etablere en selvstendig språkteknologisk industri i Norge.

Språkbanken må være nasjonens eiendom, og innholdet må være tilgjengelig for både forskning og industriutvikling. Internasjonale standarder for innsamling, merking og

tilrettelegging må anvendes for at materialet skal kunne benyttes også i flerspråklige produkt (f.eks. maskinell oversetting).

### **3.3 UTDANNING I SPRÅKTEKNOLOGI**

De mest iøynefallende utfordringene forbundet med en norsk språkteknologisatsing er at de eksisterende fagmiljøene er små og spredte, at det er forholdsvis få studenter som benytter tilbudene som gis, og at det er begrenset tilgang på digitalt lagrede språkressurser. Tallet på studenter som uteksamineres i taleteknologi, er noenlunde tilfredsstillende, mens det er en markert underproduksjon av kandidater innen datalingvistik. Det gis i 2001 undervisning i datalingvistiske emner til og med hovedfag og doktorgrad ved universitetene i Oslo, Bergen og Trondheim. Taleteknologi som undervisningsfag fins bare ved NTNU. Det gis undervisning i fonetikk ved alle landets universiteter. Universitetet i Bergen tilbyr i samarbeid med Høgskolen i Bergen grunnfag og mellomfag i språkteknologi. Faget er profesjons- og anvendelsesrettet og fyller et behov som f.eks. språkteknologimiljøet på Voss (NST) har. Dette tilbudet utfyller de øvrige tilbudene ved universitetene. Skal Norge kunne satse på språkteknologi med full tyngde, må man utdanne folk med sikte på at de skal kunne arbeide på flere nivå. Tilbudet på Voss har rekruttert studenter med ulik yrkesbakgrunn, og mange av dem kommer fra sektorer som vil anvende produktene (f.eks. helse- og sosialsektoren). De nåværende tilbud innenfor datalingvistik og språkteknologi ved utdanningsinstitusjonene bør bygges ut om man skal oppnå et tilstrekkelig rekrutteringsgrunnlag til å satse på norsk språkteknologi.

Utdanningsinstitusjonene kan selv øke antall studieplasser på et fagområde, så lenge de holder seg innenfor sine budsjettammer. Kirke-, utdannings- og forskningsdepartementet har på sin side ingen planer om å øremerke økonomiske midler for fagområdet språkteknologi. Tilbudet i Bergen er et første trinn i en profesjonsutdanning. Den bør utvides til også å omfatte hovedfag og doktorgrad og må komme i tillegg til eksisterende tilbud. Den kan f.eks. bygges opp slik det er gjort ved noen svenske universitet (Göteborg, Uppsala, Stockholm), der utdanningen omfatter både tekniske og lingvistiske emner. Universitetene har anledning til å etablere denne typen tilbud innenfor de nåværende rammene. Det er viktig at utdanningsinstitusjonene følger opp og omsetter de politiske intensjonene i handling, dersom man skal få resultater av satsingen på norsk språkteknologi.

### **3.4 FORSKNINGSPROGRAM FOR SPRÅKTEKNOLOGI**

En norsk språkbank er infrastrukturen og den nødvendige grunnmuren i en satsing på norsk språkteknologi. For å kunne reise resten av byggverket vil det være nødvendig å sette i gang forskningsprogram som kan styrke kompetansen i norsk språkteknologi og sørge for høyt kvalifiserte fagfolk. En norsk språkteknologisk satsing må ikke begrenses til applikasjonsnære prosjekt som bygger på metoder og produkt som er utviklet i andre land og for andre språk, men må omfatte fornyende grunnforskning og aktiv deltakelse i internasjonale forskningsmiljø og -fora. Språkteknologien er i så rask utvikling at det bare er gjennom aktiv deltakelse med selvstendige bidrag at norske miljø vil ha mulighet til å hevde seg.

Språkteknologien er flerfaglig av natur, og behovet for flerfaglig kompetanse er stort. En språkteknologisk satsing som sprer ressursene over små institusjoner med smalt faglig grunnlag, vil neppe lykkes. Det er først og fremst universitetene som har tilstrekkelig faglig



bredde til å utvikle levedyktige språkteknologiske FoU-miljø. De eksisterende, brede kompetansemiljøene vil derfor i samspill med språkteknologisk næringsvirksomhet ha en vesentlig rolle i utviklingen av norskspråklige produkt og tjenester på området.

Språkteknologi er i dag et internasjonalt virkefelt med både forskning og produktutvikling. Sannsynligvis verken kan eller bør Norge med sine begrensede ressurser dekke alle områder av språkteknologien. Det bør heller vurderes om det er spesielle nisjer i markedet som kan prioriteres. Innenfor anvendelsesorientert forskning og mer produktorientert virksomhet er Norge avhengig av internasjonalt samarbeid. Kvaliteten på dagens utenlandske språkteknologiprodukt som er tilpasset norsk, er ofte preget av at produsenten har hatt mangelfull norskspråklig kompetanse til rådighet, og av at tilgangen på norske språkressurser har vært begrenset.

Et språkteknologisk forskningsprogram er under utarbeiding i Norges forskningsråd. Det er viktig at programmet blir gjennomført med tilstrekkelig omfang og tyngde. Programmet er nært knyttet til planene om å etablere en norsk språkbank. Programmet skal etter planen omfatte så vel grunnforskning som anvendelsesorienterte prosjekt og stimulere til forskning og kompetanseoppbygging i prosessering av både talespråk og skriftspråk. Programplanen utarbeides sommeren 2001, og det planlegges en første utlysning av midler høsten 2001, med de første tildelingene fra 01.01.2002. Programmets faglige og tematiske avgrensning beskrives i programplanen. Det kan være aktuelt å ta med mer overordnede språkvitenskapelige og samfunnsfaglige problemstillinger som *norsk språk og IKT* og samarbeidsprosjekt med industrien for å sikre norskspråklige brukergrensesnitt i forbruksvarer (mobiltelefon, kjøleskap osv.). Alternativt kan slike prosjekt initieres av programstyret for språkteknologi og iverksettes av andre.

### **3.5 VIRKEMIDDEL – NÆRINGSVIRKSOMHET**

For å oppfylle målene om å sikre norsk språks stilling i informasjonssamfunnet og sørge for at IKT-brukere med norsk som morsmål får tilgang til produkt og tjenester på sitt eget språk, er det nødvendig å få i gang næringsvirksomhet på området. Det vil være aktuelt å utvikle og markedsføre rene språkteknologiske produkt og tjenester samt å benytte språkteknologi som en integrert del i andre produkt. Det eksisterende virkemiddelapparatet som SND, såkornfond, etablererskoler osv. har virkemiddel som er velegnet også til etablering av språkteknologisk virksomhet. I en nasjonal satsing på språkteknologi vil det være naturlig å rette en del av de offentlige virkemidlene inn mot etablering av levedyktig næringsvirksomhet på dette feltet. En norsk språkbank med nødvendige språkressurser vil også gi økte muligheter til å virkeliggjøre regjeringens mål om at IKT skal være et konkurransefortrinn for norsk næringsliv. I den grad språkteknologiske komponenter er nødvendige i nye IKT-produkt og -tjenester, vil det raskere oppnås resultater med tilgang til språkressurser som er tilrettelagt i en språkbank.

Erfaringene fra NST på Voss viser at det kan oppnås store synergieffekter gjennom å arbeide parallelt med dansk, norsk og svensk. Dette gjelder både oppbygging av språkressurser (NST har til nå nedlagt over 100 arbeidsår i ressursutvikling) og utvikling av metoder og anvendelser. På et møte arrangert av NorFA våren 2001 var det stor interesse for det norske språkbankinitiativet. Mange større virksomheter arbeider i dag på tvers av de nordiske landegrensene, og de nordiske forskningsmiljøene har god kontakt med hverandre. Forholdene burde derfor ligge vel til rette for en samordnet satsing i Norden. De ulike

særprogram under EUs rammeprogram for forskning og teknologisk utvikling bør utnyttes fullt ut, noe norske forskere og bedrifter bare i liten grad har gjort hittil. Sammenliknet med de andre nordiske landene ligger vi langt etter på dette området.

Allianser mellom forskning og industri finnes og er formalisert gjennom tiltak som SND og SIVA støtter. SND mener at kommersialisering av resultatene av språkteknologisk forskning og utvikling er en forutsetning for å sette fart på utviklingen. Et tiltak som alt er gjennomført, er etableringen av "rugekassen" SAIL Port Northern Europe AS (SPNE) på Voss. SIVA er nå eiere av SPNE, som er blitt tilført betydelige offentlige midler (totalt 43 mill. NOK til nå). SPNE har som mål å kommersialisere språkteknologiske forskningsresultat og legge forholdene til rette for nyetablering innenfor språkteknologisk virksomhet. SHD er svært opptatt av anvendelsesaspektet ved språkteknologiske produkt og tjenester og har sammen med andre satt i gang flere prosjekt rettet mot produktutvikling (IT-FUNK v/NFR, medisinsk diktering (SND+NST+RiT)).

Det bør utredes hvilken nytte og nye muligheter språkteknologiske produkt kan gi brukergrupper med særlige behov. Dette området har sannsynligvis et betydelig næringspotensial. Noe av problemstillingen er tatt opp i arbeidet som gjøres i IT-FUNK, og noen språkteknologiske produkt er i ferd med å se dagens lys som følge av satsingen på området, bl.a. lesestøtte (Multifunk) og skrivestøtte (LingDys) for personer med lese- og skrivevansker.

## Kapittel 4 NORSKSPRÅKLIG PROGRAMVARE

### MÅL

|  |
|--|
| I Norge skal IKT-produkt og -tjenester ha norskspråklig brukergrensesnitt som bygger på kunnskap om norsk tale og de to målformene   |
| Alle med norsk som morsmål skal stilles likt i tilgang til og valg av norskspråklige IKT-produkt og -tjenester og skal gis like muligheter til å skaffe seg kompetanse i bruken av dem |
| Mengden norskspråklig programvare må økes, særlig programvare på nynorsk   |
| Opplæringslovens bestemmelser om språklige parallellutgaver må oppfylles i praksis for elektroniske læremiddel   |
| I undervisningssammenheng bør elever og studenter ha tilgang til programvare på norsk (bokmål og nynorsk)  |
| En standardisert norsk IKT-terminologi må utvikles og stilles til rådighet   |
| Språket i norskspråklig programvare og på norskspråklige nettsteder må forbedres   |
| Norge skal være i fremste rekke i utvikling og bruk av nettbaserte læremiddel for skolen   |

### TILTAK – norskspråklig programvare (NP)

| Nr. | Tiltak  | Ansvarlig | Samarbeidsparter | Ref./komm.:                      |
|-----|---|-----------|------------------|----------------------------------|
| NP1 | Innføre et regelverk, med forankring i lov eller forskrift, for offentlige innkjøp av vanlig programvare (operativsystem medregnet) slik at det sikres språklige parallellutgaver | AAD/KD    |                  |                                  |
| NP2 | Gi økonomisk støtte til utvikling og bruk av verktøy for oversetting av programvare, også mellom bokmål og nynorsk  | KD        | KUF, LS          |                                  |
| NP3 | Gi økonomiske støtte til utvikling av elektroniske læremiddel på norsk (bokmål og nynorsk) og til bruk av norske språkteknologiprodukt i brukergrensesnittet deres                | KUF       | LS, KD, NHD, SND |                                  |
| NP4 | Oppheve unntaket fra kravet om språklige parallellutgaver for   | KUF       |                  | Jf. forskriftene til opplærings- |

| Nr.  | Tiltak   | Ansvarlig      | Samarbeidsparter                 | Ref./komm.:         |
|------|--|----------------|----------------------------------|---------------------|
|      | administrativ programvare til skolebruk  |                |                                  | loven, §17-3        |
| NP5  | Samle, standardisere, utvikle og gjøre tilgjengelig norsk IKT-terminologi  | Språkrådet/RTT | DND, IKT-bransjen, Nordterm, ISO |                     |
| NP6  | Utvikle og gjøre tilgjengelig språktjenester og språkkompetanse på IKT-området   | Språkrådet     | DND, IKT-bransjen                |                     |
| NP7  | Informere leverandører og kjøpere om regelverket for språklige parallellutgaver av IKT-produkt                         | KUF            |                                  |                     |
| NP8  | Lage og gjøre tilgjengelig en oversikt over elektroniske læremiddel som oppfyller kravet om språklige parallellutgaver | KUF            | LS                               | Noe er gjort hos LS |
| NP9  | Initiere seminar og konferanser der språk og IKT er tema   | Språkrådet     | KD, NHD, SND                     |                     |
| NP10 | Etablere et nettsted med informasjon om og formidling av språkressurser og språktjenester                              | Språkrådet     | KD, NHD, SND, IKT-bransjen       |                     |
| NP11 | Opprette språkpris e.l. for programvare, elektroniske læremiddel og nettsteder   | Språkrådet     | DND, LS, IKT-bransjen            | Jf. Rosingprisen    |

I dette kapitlet vil de mest vidtfavnende emnene omtales først, deretter de mer sektorrelaterte. Avsnittet om IKT-terminologi kommer først fordi terminologien gjennomsyrrer alle typer programvare. Bruk av en standardisert norsk terminologi vil bidra til å heve den språklige kvaliteten i all slags programvare.

## 4.1 NORSK IKT-TERMINOLOGI

Godt fagspråk medvirker til rask og effektiv kommunikasjon fagfolk imellom. For den vanlige bruker er det viktig at terminologien som anvendes i IKT-produkt, er mest mulig selvforklarende, og at den blir brukt konsekvent. Hvilket språk som dominerer en gitt fagterminologi, bestemmes i all hovedsak av i hvilket eller hvilke land fagutviklingen foregår raskest og i størst omfang. Innenfor IKT skjer dette i dag i engelsktalende land, og IKT-terminologien er derfor kraftig preget av engelsk. På IKT-området er engelsk innarbeidet som felles språk for utveksling av informasjon mellom eksperter og for presentasjon av nye produkt for et marked. Mye av den engelskspråklige terminologien trekkes ukritisk over i norsk allmennspråk uten tilpasning. I flere språkfaglige miljø arbeides det med å finne egnede

norske avløserord. Noen ganger godtas engelskspråklige termer med norsk skrivemåte og bøyning ("norvagisering"), og iblant anvendes de uten fornorsking.

Kompetanse i terminologiarbeid utvikles best i krysningpunktet mellom språkkompetanse og kompetanse i det faget termene tilhører. Arbeidet må drives kontinuerlig og systematisk for å kunne være til hjelp for dem som har behov for termene. IKT-bransjen må bidra med de nødvendige fagkunnskaper i dette arbeidet. I forbindelse med utvikling av verktøy for oversetting og lokalisering (kulturell tilpasning) av programvare vil tilgang på relevante terminologiresurser være avgjørende for å lette og effektivisere arbeidet. Terminologien bør fortrinnsvis være integrert i verktøyet, men vil også ellers være til stor nytte under arbeidet.

En norsk IKT-termbase bør i tillegg til norske oppslagsord og ekvivalenter på andre språk inneholde definisjoner og/eller forklaringer. Mest aktuelt er det å ha med termer for de andre nordiske språkene og engelsk, fransk og tysk. Det er etablert et overordnet terminologisamarbeid på både nordisk og internasjonalt plan der man standardiserer terminologi på flere områder, hvorav IKT er ett. På nordisk plan koordineres terminologiarbeidet i samarbeidsorganisasjonen Nordterm, som jevnlig holder møter og i tillegg arrangerer kurs i terminologirelaterte emner. ISO er en internasjonal standardiseringsorganisasjon der IKT-terminologi er et av arbeidsområdene. Nordterm følger ISO-standardene, og på denne måten koples bl.a. det norske standardiseringsarbeidet til det internasjonale. ISO-standardene vil være til stor hjelp for norske programutviklere og oversettere når de skal finne fram til riktig norsk term.

Det er behov for å gjennomgå og avklare ansvarsfordelingen mellom Språkrådet og Rådet for tekniske terminologi (RTT) når det gjelder hovedansvaret for arbeidet med norsk IKT-terminologi.

## 4.2 SPRÅKET I PROGRAMVARE

Tilbud av norskspråklige produkt vil være en viktig forutsetning for å nå målet om at alle skal få lik tilgang til kunnskapssamfunnet (jf. *Plan for styrking av norsk*, Norsk språkråd 2001). Utgangspunktet er en bevisst holdning til hvilken språkutvikling vi ønsker i Norge. Tilgang til språkressurser og -tjenester på Internett kan gjøre det enklere for produsentene å høyne kvaliteten på språket i produktene. Det bør være et konkurransefortrinn å kunne tilby produkt på begge målformer og med høy språklig kvalitet.

I dag møter de fleste sluttbrukere, uansett alder og yrke, et grafisk grensesnitt på elektroniske innretninger. På datamaskinen er dette gjerne MS Windows. Mye norskprodusert programvare bygger på utenlandske verktøy eller applikasjoner. Utviklingsspråket for norske produsenter vil ofte være engelsk fordi norsk programvare må være kompatibel med annen programvare. I et marked der neste versjon er avgjørende for videre suksess, er kvalitetssikring av språket blitt en nedprioritert oppgave i forhold til å sørge for teknisk velfungerende produkt. Det er viktig at brukerne møter et norskspråklig grensesnitt (bokmål eller nynorsk) fra de tar i bruk teknologien for første gang. Blant eldre kan et norskspråklig grensesnitt redusere frykten for teknologien og senke terskelen for å ta den i bruk.

Over 4 000 norske bedrifter lever av å produsere, importere og/eller distribuere programvare for det norske markedet. Over 1 000 av dem tilbyr egenprodusert programvare i en eller annen form (kilde: IKT-Norge). Spillprodusentene, som i hovedsak sikter seg inn mot

eksportmarkedet, lager norskspråklige versjoner i tillegg til fremmedspråklige. For slike programvareprodusenter med ambisjoner også på det internasjonale markedet vil hensiktsmessige språkversjoneringsverktøy være et viktig hjelpemiddel. Utenlandske aktører bør oppmuntres til å lage norskspråklige versjoner av sine nettsteder. Det vil sannsynligvis være enklere å påvirke dem i ønsket retning dersom oversettings-/versjoneringsverktøy med terminologiresurser står til disposisjon.

Det offentlige må kunne stille krav om parallelle programutgaver på bokmål og nynorsk ved innhenting av tilbud i forbindelse med innkjøp. Verktøy som kan støtte og effektivisere arbeidet med oversetting mellom de to målformene, vil med stor sannsynlighet bli godt mottatt av IKT-bransjen. Det finnes allerede et norskutviklet program for oversetting fra bokmål til nynorsk (Nyno). Det finnes også generelle verktøy for oversetting mellom språk (bl.a. Trados og IBMs Translation Manager). Slike verktøy vil også kunne benyttes til oversetting mellom bokmål og nynorsk. Det vil være lettere for IKT-bransjen å tilby norskspråklige produkt på begge målformer dersom offentlige myndigheter påtar seg noe av utgiftene til videreutvikling og tilrettelegging for bruk av slike verktøy.

### **4.3 KONTORSTØTTEPROGRAM OG ADMINISTRATIV PROGRAMVARE**

Kontorstøtteprogram er de vanligste programmene for storparten av brukerne. Kategorien omfatter program for tekstbehandling, regneark, elektronisk post, tegneprogram, presentasjonsprogram, enkle databaseprogram osv. og markedsføres gjerne i pakker med varierende innhold. Slike programpakker må tilbys med brukergrensesnitt, dvs. menyer, ledetekster og hjelpefiler, på begge de norske målformene. Standardmaler som følger programvaren, må også tilbys på bokmål og nynorsk. Velfungerende tilleggsprogram for orddeling, stavekontroll, grammatikkontroll o.l. og andre skrivestøtteprogram må foreligge til samme tid for begge målformer. Dette vil bidra til å sikre norsk språkkultur.

Administrativ programvare omfatter saksbehandlerverktoy, kundeinformasjonssystem, økonomisystem, regnskapssystem, gruppevare, bransjesystem, offentlige registre og andre system som benyttes i administrativ sammenheng i en virksomhet. Systemene har gjerne en database i botnen. Saksbehandlere og andre sluttbrukere må kunne møte sin foretrukne målform når de bruker denne typen programvare.

I skoleverket er kontorstøtteprogram og administrativ programvare ofte omtalt samlet som "administrativ programvare". I den grad elever i grunnskolen og videregående opplæring bruker slike program i skolesammenheng, må de være tilgjengelige på begge målformer.

I målloven stilles det krav til statsorgan om at alle skal få svar på henvendelser og kunne få ferdige skjema o.l. på den målformen de selv bruker. Det betyr at alle administrative saksbehandlingsverktøy må ha mulighet for lagring og gjenfinning av informasjon på begge målformer.

### **4.4 ELEKTRONISKE LÆREMIDDEL**

Informasjons- og kommunikasjonsteknologien vil forandre utdanningssystemet og de pedagogiske arbeidsmetodene. KUFs satsing på IKT gjelder hele utdanningssystemet fra

grunnskole til høyere utdanning (medregnet voksenopplæring og etter- og videreutdanning) og omfatter både de faglige, pedagogiske og organisatoriske sidene ved systemet. I KUFs handlingsplan *IKT i norsk utdanning. Plan for 2000-2003* er følgende målformulering benyttet: ”*IKT i utdanningen skal bidra organisatorisk, faglig og pedagogisk til et utdanningssystem som utvikler og utnytter IKT som fag og som utnytter fullt ut de muligheter IKT gir i undervisning og læring, slik at den enkeltes og samfunnets kompetansebehov imøtekommes*”(s. 6).

Utdanningssektoren har i dag tatt i bruk den nye teknologien på mange områder, men tilgang på utstyr, nettilknytningsmåte og -grad samt kompetansenivå varierer. Barn fra førskolealder og oppover er fortrolige med teknologien og bruker med største selvfølgelighet engelskspråklige data- og tv-spill. Skolen oppfordrer elevene til å søke etter informasjon på Internett. Holdninger skapes tidlig, og ønsker vi at norsk fortsatt skal være et fullgodt bruksspråk i alle sammenhenger, må det gjøres et systematisk framstøt overfor produsentene av IKT-produkt rettet mot barn og unge (Jf. *Plan for styrking av norsk*, Språkrådet 2001).

Elever i grunnskolen og videregående opplæring er godt kjent med og bruker i varierende grad det som i kapittel 4.3 er omtalt som kontorstøtteprogram. I §17-3 i forskriftene til opplæringsloven er denne typen programvare omtalt som administrativ programvare og unntatt fra kravet om språklige parallellutgaver. Parallellutgaver av vanlige brukerprogram er et generelt krav i denne handlingsplanen, og elevenes utstrakte bruk av slike program i skolesammenheng tilsier at forskriftene må endres på dette punktet. Stortingets kirke- og utdanningskomité har bedt KUF legge fram en tidsplan for oppheving av unntaket. LS utreder nå saken.

I dette avsnittet brukes ”elektronisk læremiddel” om alt elektronisk materiale som tilfredsstillende definisjonen av læremiddel i §17-2 annet og tredje ledd i forskriftene til opplæringsloven, og som brukes i undervisningssammenheng. Pedagogisk programvare faller inn under elektroniske læremiddel og er opplæringsprogram som er tilrettelagt for interaktiv bruk i undervisningen i et fag. Elektroniske utgaver av trykte lærebøker som ikke er spesielt tilrettelagt for interaktiv bruk, er elektroniske læremiddel, men ikke pedagogisk programvare. Det samme gjelder materiale som er lastet ned fra Internett i undervisnings- og/eller læringsøyemed, men som ikke er spesielt tilrettelagt for undervisningsformål. Pedagogisk programvare er med andre ord kjennetegnet ved interaktivitet mellom brukeren og programvaren. KUF har gjort det helt klart at kravet om språklige parallellutgaver omfatter elektroniske læremidler (*IKT i norsk utdanning – årsplan for 2001* (KUF, mai 2001)). Aktører som legger ut læringsressurser på nettsteder, bør derfor påse at de tilbyr stoff på begge målformer når forskriftene til opplæringsloven krever det.

For å sikre at bestemmelsene om språklige parallellutgaver i §17-1 til §17-4 i forskriftene til opplæringsloven etterleves også med hensyn til elektroniske læremiddel, må KUF informere produsentene og skoleeierne skriftlig om de regler som gjelder. Samtidig bør man understreke det ansvaret skoleeierne har etter forskriftenes § 17-5. En god hjelp for skolene i så måte er den oversikten over språklige parallellutgaver som LS ajourfører, og som er lagt ut og gjort tilgjengelig på nettstedet Skolenettet. Oversikten bør utvides til å omfatte elektroniske læremiddel.

Det trengs egne støtteprogram for elever og studenter med særlige behov som følge av lærevansker eller funksjonshemninger. Språkteknologiske verktøy vil være til særlig stor nytte for denne målgruppen (jf. eNorge-planen, versjon 2.0 pkt. 2.1.3).

Program til bruk i norskopplæringen for fremmedspråklige bør være flerfunksjonelle og gi mulighet for lesing, lytting, innlesing og sammenligning av uttale samt besvaring av oppgaver. I denne typen program bør det være integrert språkteknologiske produkt som kan gi målgruppen hjelp med hensyn til synonymmer, grammatikk, semantikk, stilistikk, setningsoppbygging og tekstanalyse.

I allmennlærerutdanningen tilbys det i dag en del kurs i IKT, og de fleste fag har noe anvendelsesorientert opplæring i bruk av IKT som støtte i undervisningen. På universitetsnivå har den generelle kapasiteten innen IKT-utdanningen økt betraktelig de seinere årene. Det er også økende interesse for bruk av IKT som verktøy for tilrettelegging av lærestoff på Internett. Nettbasert undervisning, ofte omtalt som e-læring, har et stort utviklingspotensial. I framtiden vil det være vanlig at enkeltpersoner og bedrifter benytter slike tilbud i forbindelse med etterutdanning og andre kompetansehevingstiltak. Kvalitetskravene til disse tjenestene må omfatte språklige aspekter.

Noen miljø ved norske universitet og høyskoler har arbeidet med forskning i og utvikling av pedagogiske metoder for nettbasert undervisning. Det foregår dessuten utprøving av ulike typer teknologi for pedagogiske formål (interaktive læringsmiljø på Internett). Det finnes internasjonale tilbud om opplæring på Internett på grunnlag av tilpasset norskutviklet programvare og kursinnhold. Tilbudene gjelder både IKT-faget og andre fag. Kompetansen i nettbasert undervisning må utnyttes og bygges videre på i utviklingen av norskspråklige elektroniske læremiddel. Norske bedrifter har gode muligheter til å etablere seg som produsenter av programvare til slik bruk og som leverandører av innhold og tekniske løsninger.

## **4.5 TEKNISK PROGRAMVARE**

Profesjonsrelatert teknisk programvare er utviklet for spesielle bruksformål. Denne typen programvare har et forholdsvis smalt internasjonalt marked, og det er ikke hensiktsmessig å stille spesielle språklige krav til den. Man bør likevel arbeide for at rapporter og utskrifter for allmenne formål genereres på norsk i Norge.

Teknisk programvare for vanlige brukere benyttes til f.eks. drift og overvåking av store, komplekse system (energiforsyning, industriproduksjon, databasedrift o.a.). For å sikre at operatørene kan utføre sitt arbeid effektivt, nøyaktig og uten misforståelser med alvorlige følger, bør slike system ha meldinger, alarmer og andre tekster på norsk.

En spesiell type teknisk programvare for alminnelige brukere blir stadig vanligere i forskjellige nyttegenstander i hjemmet. Kjøleskap, komfyr, fryseboks, panelovner o.l. leveres med programvare som gjør at de kan styres direkte fra en meny, med telefon eller over Internett. Det har lenge vært reklamert med at "du kan ringe opp hytta slik at den er varm til du kommer". Forutsetningen er at man har fått installert en prosessor som styrer de elektriske panelovnene. Dagligvarebutikker vil få selvbetjente kasser. For at kundene skal kunne gjøre bruk av dem og kontrollere kassalappen, må programmene i kassasystemet ha norske tekster. Denne typen programvare vil sannsynligvis få elektronisk bruksanvisning i stedet for papirbasert, og etter hvert vil den elektroniske bruksanvisningen bli avløst av kunstig tale (talesyntese). Dette er et område hvor det åpenbart er behov for norskspråklige menyer og bruksanvisninger, enten de er elektroniske eller utnytter kunstig tale.



Det bør utredes nærmere om grensesnittspråket i enkelte av produktene nevnt ovenfor kommer inn under bestemmelsene i arbeidsmiljølovens (AML) kapittel IV, som omhandler leverandørers ansvar for menyer og bruksanvisninger. Regjeringen vil ta for seg dette problemfeltet i forbindelse med arbeidet med en ny arbeidslivslov (jf. eNorge-planen, versjon 2.0 pkt. 2.4.1). Det bør vurderes innført et krav om norsk språk i elektroniske menyer og bruksanvisninger, fordi språkvalget kan være av betydning for arbeidsmiljøet og ha et viktig sikkerhetsmessig aspekt.

## 4.6 NETTSTEDER OG UNDERHOLDNINGSVARER

Mange nettsteder skjemmes av dårlig norsk. Mye kan gjøres for å heve kvaliteten på språkbruken, og det bør være et mål å gjøre godt norsk språk til et kvalitetsstempel som eiere av og bidragsytere til nettsteder kan være stolte av. En standard utforming vil gi økt funksjonalitet og lesbarhet. World Wide Web Consortium (W3C) har etablert et "Web Accessibility Initiative" med retningslinjer for utforming av nettsider, slik at brukerne kan få optimalt utbytte av sidene (<<http://www.w3c.org/WAI>>). Norske utviklere av nettsider kan med fordel benytte retningslinjene.

Bredbånd vurderes som et viktig grunnlag for næringsutvikling i Norge (jf. regjeringens handlingsplan for bredbåndskommunikasjon, NHD 2000). Bibliotek og mediebedrifter har store mengder materiale på norsk som kan egne seg som innhold i bredbåndstjenester for vanlig Internett, Internett på mobiltelefon og interaktivt fjernsyn. Ser vi framover, spør ekspertene at følgende tjenestetyper vil dominere datatrafikken i de norske mobilnettene:

- Underholdningsspill
- E-post
- Bildepostkort
- Lokasjonstjenester
- Betalingstjenester og e-handel
- Interaktiv rådgivning
- Pengespill

Startsidene ("portalene") for tjenestene vil bli viktige fordi de setter rammene for brukernes valg av tjenester. De som har kontroll over startsidene, får viktige maktposisjoner i markedet. Bredbåndsstartsidene for interaktivt fjernsyn vil inneholde mye underholdning i form av video og dataspill. Internett-startsidene vil først og fremst ha en nyttefunksjon i jobbsammenheng. Startsidene for Internett på mobiltelefon vil være sterkt preget av personlig tilpasning.

Det er viktig at alle som bruker og dermed forvalter norsk språk, er klar over denne utviklingen og forholder seg til de utfordringene den bringer. Det bør kartlegges hvordan vi kan få mest mulig norsk innhold i de nye tjenestene, og hva som eventuelt hindrer IKT-bransjens tilgang til norsk innhold.

## 4.7 NORSK E-HANDEL

Det nyopprettede nettstedet <handel.no> har som målgruppe alle som er eller vil bli berørt av elektronisk handel. Nettstedet skal være en læringsarena for bedrifter som har begynt eller planlegger å begynne med elektronisk handel, og et felles møtested for slike e-handelsbedrifter og tilbydere av informasjon, kompetanse og løsninger på området (jf. regjeringens pressemelding i forbindelse med lanseringen av nettstedet). Det er ønskelig at nettstedet får lenker til relevante ordlister, terminologilister o.l. til hjelp for både tilbydere og e-handelsbedrifter. Det burde også inneholde opplysninger om tilgjengelige språktjenester og lenker til nettsteder med slike tjenester.

Statens informasjonstjeneste har laget en veiledning for hvordan det offentlige bør presentere seg selv og sine tjenester på Internett. Det bør være mulig å utvikle en tilsvarende veiledning for det private næringsliv (jf. omtalen av W3C i kap. 4.6). I en slik sammenheng er det naturlig at språk og språkbruk gjøres til et vesentlig tema, ikke minst fordi klart og entydig språk bidrar til å hindre misforståelser hos alle innblandede parter.

## Kapittel 5 SAMARBEID OG NETTVERKSBYGGING

### MÅL

Norge skal bidra til å styrke erfaringsutvekslingen mellom språkorganisasjonene i Norden og ellers i Europa om informasjons- og kommunikasjonsteknologiens innvirkning på nasjonalspråkene og om tiltak for å holde språkene i hevd også på dette området

Norge skal delta aktivt i det nordiske og øvrige europeiske samarbeidet innenfor språkteknologi

Norge skal delta aktivt i det nordiske og internasjonale arbeidet med utvikling og standardisering av IKT-terminologi

### TILTAK – samarbeid og nettverksbygging (SN):

| Nr. | Tiltak   | Ansvarlig  | Samarbeidsparter               | Ref./komm.:                   |
|-----|--|------------|--------------------------------|-------------------------------|
| SN1 | Utrede behovet for en sektorovergripende enhet som kan koordinere nasjonale tiltak innenfor språk og IKT             | Språkrådet | KD, NHD, KUF, NFR              | Jf. St.meld. nr. 13 (1997-98) |
| SN2 | Ta initiativ til at det opprettes nordiske samarbeidsfora eller -enheter for språkteknologi                          | Språkrådet | KD, NFR, NHD, KUF, Nordisk råd |                               |
| SN3 | Initiere, støtte og delta aktivt i tiltak knyttet til språk og IKT (på nasjonalt, nordisk og øvrig europeisk plan)   | Språkrådet | KD, NFR, KUF, Nordisk råd, EU  |                               |
| SN4 | Støtte norske aktører økonomisk og praktisk slik at de blir i stand til å lede og delta i prosjekt på europeisk plan | NFR        | KUF, universitetene            |                               |

Verken i Norge eller Norden for øvrig finnes det noen sektorovergripende og handlekraftig enhet eller instans som koordinerer tiltak på området språk og IKT. Internett bidrar til at kommunikasjonsmedier som tidligere var atskilte, smelter sammen og oppfattes som ett. Internett gir tilgang til tekst, bilder, animasjoner, video, 3D-modeller og lyd. Ett og samme grensesnitt gir mulighet til å se TV, skrive brev, lage presentasjoner og selv kople lyd, bilde og tekst. Bare et fåtall kjenner til de muligheter som her foreligger, og kan utnytte dem fullt ut. Skal man oppnå økt bruk av IKT-tjenester og utvide den enkeltes tilgang til dem, er språket viktig. Nasjonalt er det mye å vinne med en sektorovergripende koordinering av innhold som tilrettelegges og gjøres tilgjengelig på Internett. Bredbåndssatsingen er en gyllen anledning til å prøve ut hvor mye man kan oppnå ved en koordinert gjennomføring av strategier. Koordinering betyr bedre utnytting av ressursene for både tilbyder og bruker, og mer gjenbruk av innholdet gjør tjenestene billigere.

Internasjonalt har forskere alltid samarbeidet på sine respektive fagfelt, utvekslet erfaringer og drevet prosjekt sammen. Hvis Norge blir blant de fremste i utviklingen av språkteknologiske

verktøy, ren teknologi og/eller oppbygging av en bank med gjenbrukbare språkressurser, vil det være et viktig bidrag til innsatsen for å bevare og styrke det språklige mangfoldet i Europa (jf. EUs eContent-program og eEurope-plan). Flerspråklige korpus vil være til stor nytte for alle parter ved samarbeid på tvers av landegrensene og vil dessuten være et viktig materiale for forskere og utviklere i industrien som arbeider med oversettingsverktøy og elektroniske taletjenester for flere språk.

Forskningsmiljøene har i varierende grad formelle og uformelle nettverk for samarbeidet. Studentutveksling er en viktig del av det mer formelle samarbeidet på faglig og organisatorisk plan. Det finnes interessante studieprogram også i andre nordiske land, bl.a. ved Kungliga Tekniska Högskolan i Stockholm, Universitetet i Göteborg og Chalmers Tekniska Högskola i samme by. Disse programmene kan brukes systematisk i oppbyggingen av et felles nordisk språkteknologimiljø. Den enkelte forsker gjør vanligvis selv avtaler om gjesteopphold ved en utenlandsk institusjon, men forholdene kan tilrettelegges bedre på dette området, særlig på nordisk plan. Nordisk Forskerutdanningsakademi (NorFA) har utredet mulighetene for et samarbeid om en felles forskerutdanning i Norden. På De nordiske datalingvistikkdagene 2001 (Uppsala, mai 2001) diskuterte man muligheten for en felles doktorgradsutdanning i språkteknologi og konkluderte med at man burde fortsette sonderingene på grunnlag av diskusjonen. I tillegg bør det utredes om det kan være nyttig med en egen enhet som koordinerer forskning og produktutvikling innenfor språk og IKT på nordisk plan. Nordisk Ministerråd har bevilget i alt 23 mill. DKK til et språkteknologisk forskningsprogram som startet i 2000 og løper fram til 2004. Åtte prosjekt er tildelt midler, deriblant tre ledet av norske forskere. Programmet administreres av NorFA. Nordisk språkråd har tatt initiativ til å opprette et kontaktnett for IKT-spørsmål og IKT-terminologi for de nordiske språknemndene.

Innenfor språkteknologisk forskning og produktutvikling kan et nordisk samarbeid være kostnadseffektivt og utbytterikt for både forskere og næringsutviklere. Man vil kunne dra nytte av at dansk, norsk og svensk har mange likhetstrekk. Produkt framstilt for et av språkene vil kunne tilpasses de andre med lavere omkostninger enn tilfellet er med engelskspråklige produkt. Kulturen i de nordiske landene har mange fellestrekk, og det bør derfor være mulig å utnytte hverandres nettsteder.

Deltakelse i EU-støttede forskningsprosjekt forutsetter minst 50 % finansiering fra andre kilder. I mange tilfeller kan dette hindre enkeltforskere, grupper av forskere eller institusjoner i å delta. Det bør bli enklere for norske forskere å delta i disse forskningsprosjektene og å påta seg å lede dem. En løsning kan være å øremerke nasjonale midler for slik deltakelse.

## Kapittel 6 KONKLUSJONER

### 6.1 HANDLINGSPLANEN

I denne handlingsplanen forutsettes det at det finnes politisk vilje til å virkeliggjøre visjonen om at informasjons- og kommunikasjonsteknologien skal bli tilgjengelig for alle i Norge. Det forutsettes videre at norsk skal være det foretrukne språket i Norge også i IKT-sammenheng. Språkrådets plan for styrking av det norske språket motiverer til innsats på flere områder for å forebygge at norsk trues som bruks- og fagspråk. Handlingsplanen for norsk språk og IKT støtter opp om dette ved at det der pekes på viktige språkstyrkingstiltak som kan forebygge uheldige følger av informasjonssamfunnet. Eksempel på dette er etablering av en norsk språkbank, igangsetting av et språkteknologisk forskningsprogram og tettere, aktivt samarbeid mellom kultursektoren og den private sektor om tiltak for å forbedre språket i eksisterende og nye IKT-produkt.

Det foreslåtte tiltaket for å samle og gjøre tilgjengelig norske språkressurser i en språkbank er det viktigste, mest omfattende og mest krevende i planen, målt i både arbeidsinnsats og kostnader. Dette tiltaket er også det mest presserende på bakgrunn av de påtrengende utfordringene det norske språksamfunnet står overfor i tilknytning til bruk av språk i moderne teknologi.

Oversetting av programvare fra bokmål til nynorsk har høy prioritet i Språkrådet. Dette er arbeid som må settes i gang umiddelbart. Kostnadene ved arbeidet er avhengig av hvor mange program som blir oversatt, hvor mye som er likt i de oversatte programmene, hvem som utfører arbeidet, og hvorvidt det blir utført ved hjelp av tilrettelagte oversettingsverktøy eller ikke. Utvikling av slike verktøy for språkversjonering er også høyt prioritert, først og fremst for bruk mellom bokmål og nynorsk, men også for oversetting mellom norsk og fremmedspråk. Kostnadene forbundet med dette må utredes. Et samarbeidsprosjekt med kostnadsdeling mellom det offentlige og sentrale aktører i IKT-bransjen vil forhåpentlig raskt gi resultater. Oversetting av vanlige kontorstøtteprogram og program som brukes i skolen, må ha høyeste prioritet.

Enkelte tiltak som er foreslått i planen, kan gjennomføres uavhengig av hverandre. Noen tiltak forutsetter bedre utnytting av mulighetene i eksisterende teknologi. Noen peker på muligheter for utvikling av og tilgang til nye tjenester som bygger på dagens teknologi, mens andre forutsetter utvikling av helt ny teknologi og en ny type tjenester.

### 6.2 TILTAK ETTER KAPITTEL

#### Kapittel 3: TILTAK – norsk språkteknologi (SP):

| Nr. | Tiltak  | Ansvarlig | Samarbeidsparter                  | Ref./komm.      |
|-----|---|-----------|-----------------------------------|-----------------|
| SP1 | Etablere en norsk språkbank                               | KD        | Språkrådet, NFR, NHD              | Utredes i 2001  |
| SP2 | Sørge for et fullverdig utdanningstilbud i språkteknologi | KUF       | Universitets- og høyskolesektoren |                 |
| SP3 | Opprette et forskningsprogram i språkteknologi            | NFR       | KUF, NHD, KD                      | Er under arbeid |

| Nr. | Tiltak   | Ansvarlig | Samarbeidsparter                   | Ref./komm. |
|-----|--|-----------|------------------------------------|------------|
| SP4 | Fremme industriell utnytting av norsk språkteknologi   | NHD       | NFR, SND, SIVA                     |            |
| SP5 | Utrede behov for og legge til rette språkteknologiske produkt og tjenester beregnet på brukere med særlige behov | SHD       | KUF, NHD, NFR (IT-FUNK, TYIN), AAD |            |

#### Kapittel 4: TILTAK – norskspråklig programvare (NP):

| Nr. | Tiltak  | Ansv.          | Samarbeidsparter                 | Ref./komm.:                                 |
|-----|---|----------------|----------------------------------|---|
| NP1 | Innføre et regelverk, med forankring i lov eller forskrift, for offentlige innkjøp av vanlig programvare (operativsystem medregnet) slik at det sikres språklige parallellutgaver | AAD/KD         |                                  |   |
| NP2 | Gi økonomisk støtte til utvikling og bruk av verktøy for oversetting av programvare, også mellom bokmål og nynorsk  | KD             | KUF, LS                          |   |
| NP3 | Gi økonomiske støtte til utvikling av elektroniske læremiddel på norsk (bokmål og nynorsk) og til bruk av norske språkteknologiprodukt i brukergrensesnittet deres                | KUF            | LS, KD, NHD, SND                 |   |
| NP4 | Oppheve unntaket fra kravet om språklige parallellutgaver for administrativ programvare til skolebruk   | KUF            |                                  | Jf. forskriftene til opplæringsloven, §17-3 |
| NP5 | Samle, standardisere, utvikle og gjøre tilgjengelig norsk IKT-terminologi   | Språkrådet/RTT | DND, IKT-bransjen, Nordterm, ISO |   |
| NP6 | Utvikle og gjøre tilgjengelig språktjenester og språkkompetanse på IKT-området  | Språkrådet     | DND, IKT-bransjen                |   |
| NP7 | Informere leverandører og kjøpere om regelverket  | KUF            |                                  |   |

| Nr.  | Tiltak   | Ansv.      | Samarbeidsparter           | Ref./komm.:            |
|------|--|------------|----------------------------|------------------------|
|      | for språklige parallell-<br>utgaver av IKT-produkt   |            |                            |                        |
| NP8  | Lage og gjøre tilgjengelig<br>en oversikt over<br>elektroniske læremiddel<br>som oppfyller kravet om<br>språklige parallellutgaver | KUF        | LS                         | Noe er gjort<br>hos LS |
| NP9  | Initiere seminar og<br>konferanser der språk og<br>IKT er tema   | Språkrådet | KD, NHD, SND               |                        |
| NP10 | Etablere et nettsted med<br>informasjon om og<br>formidling av språkres-<br>surser og språktjenester                               | Språkrådet | KD, NHD, SND, IKT-bransjen |                        |
| NP11 | Opprette språkpris e.l. for<br>programvare, elektroniske<br>læremiddel og nettsteder   | Språkrådet | DND, LS, IKT-bransjen      | Jf. Rosing-<br>prisen  |

#### **Kapittel 5: TILTAK – samarbeid og nettverksbygging (SN):**

| Nr. | Tiltak  | Ansv.      | Samarbeidsparter                 | Ref./komm.:                         |
|-----|---|------------|----------------------------------|-------------------------------------|
| SN1 | Utrede behovet for en<br>sektorovergripende enhet<br>som kan koordinere<br>nasjonale tiltak innenfor<br>språk og IKT            | Språkrådet | KD, NHD, KUF, NFR                | Jf. St.meld.<br>nr. 13<br>(1997-98) |
| SN2 | Ta initiativ til at det<br>opprettas nordiske<br>samarbeidsfora eller<br>-enheter for språkteknologi                            | Språkrådet | KD,NFR, NHD, KUF,<br>Nordisk råd |                                     |
| SN3 | Initiere, støtte og delta aktivt<br>i tiltak knyttet til språk og<br>IKT (på nasjonalt, nordisk<br>og øvrig europeisk plan)     | Språkrådet | KD, NFR, KUF, Nordisk<br>råd, EU |                                     |
| SN4 | Støtte norske aktører økono-<br>misk og praktisk slik at de<br>blir i stand til å lede og delta<br>i prosjekt på europeisk plan | NFR        | KUF, universitetene              |                                     |

## FORKORTELSER BRUKT I HANDLINGSPLANEN

|           |   |
|-----------|---|
| AAD       | Arbeids- og administrasjonsdepartementet          |
| AML       | Arbeidsmiljøloven                                 |
| DND       | Den norske dataforening                           |
| FoU       | Forskning og utvikling                            |
| IKT       | Informasjons- og kommunikasjonsteknologi          |
| IKT-Norge | Medlemsorganisasjon for IKT-bedrifter i Norge     |
| ISO       | Den internasjonale standardiseringsorganisasjonen |
| IT        | Informasjonsteknologi                             |
| KD        | Kulturdepartementet                               |
| KRD       | Kommunal- og regionaldepartementet                |
| KUF       | Kirke-, utdannings- og forskningsdepartementet    |
| LS        | Læringscenteret, enhet underlagt KUF              |
| MLIS      | Multi-Lingual Information Society (se neste side) |
| ODIN      | Offentlig dokumentasjon og informasjon i Norge    |
| NFR       | Norges forskningsråd                              |
| NHD       | Nærings- og handelsdepartementet                  |
| NorFA     | Nordisk Forskerutdanningsakademi                  |
| NTNU      | Norges teknisk-naturvitenskapelige universitet    |
| RTT       | Rådet for teknisk terminologi                     |
| SD        | Samferdselsdepartementet                          |
| SHD       | Sosial- og helsedepartementet                     |
| SIVA      | Selskapet for industrivekst                       |
| SND       | Statens nærings- og distriktsutviklingsfond       |



## AKTUELLE DOKUMENT/UNDERLAG FOR HANDLINGSPLANEN

Arbeidsmiljøloven

*Den norske IT-veien. Bit for bit, rapport*, Statssekretærutvalget for IT, SD 1996

eContent Rådsvedtak av 22. desember 2000 om vedtakelse av et flerårig fellesskapsprogram for å stimulere til utvikling og bruk av europeisk digitalt innhold i verdensomspennende

*nett og for å fremme språklig mangfold i informasjonssamfunnet* (301D0048, uoffisiell norsk oversettelse) – EU-program for 2001-2004 (avløser og viderefører MLIS og INFO2000)

eEurope *An Information Society for All* – EU-kommisjonens handlingsplan, 1999 og senere (motsvares av eNorge)

eNorge, handlingsplan, versjon 1.0 (06/2000), 2.0 (12/2000) og midtveisrapport 09.03.2001, NHD (den norske regjeringens motstykke til eEurope)

*Er norsk et naturlig språk?* Victoria Rosén i Ø. Andersen m.fl. 2000

*Forskerutdannelse i Norden*, NorFA 2001

*Forskrift til opplæringslova*, Rundskriv 40, KUF 1999

*IKT i norsk utdanning. Plan for 2000-2003*, KUF 2000

INFO2000 Rådsvedtak av 20. mai 1996 om vedtakelse av et flerårig fellesskapsprogram som skal stimulere utviklingen av en europeisk multimedie-innholdsindustri og fremme bruken av multimedie-innhold i det framvoksende informasjonssamfunnet (396D0339) – EU-program for 1996-2000 (videreført gjennom eContent)

Innst.S.nr. 174 (1997-98) *Innstilling frå familie-, kultur- og administrasjonskomiteen om målbruk i offentleg teneste*

Lov om målbruk i offentleg teneste, 1981

MLIS Rådsvedtak av 21. november 1996 om vedtaking av eit fleirårig program for å fremje språkmangfaldet i Fellesskapet i informasjonssamfunnet (396D0664) – EU-program for 1995-99 (videreført gjennom eContent)

*Nasjonalt program for elektronisk handel og forretningsdrift med fokus på SMB*, programbeskrivelse, NHD 2000

*Norge – en utkant i forkant. Næringsrettet IT-plan 1998-2001*, NHD februar 1998

*Norsk språkbank. Utredning om et nasjonalt korpus for språkteknologi*, Svendsen o.a. 1999

*Plan for styrking av norsk*, Norsk språkråd 2001

*Regjeringens handlingsplan for bredbåndskommunikasjon*, NHD 2000

*Regjeringens program for fornyelse av offentlig sektor*, AAD 2000

*Satsing på informasjonsteknologi for funksjonshemmede (IT-FUNK)*, 1998-2001, NFR 1998

*"Si@!". Elektronisk samhandling i helse- og sosialsektoren, statlig tiltaksplan 2001-2003*, SHD 2001

*Skape – bevare – formidle. Handlingsplan for IT på kulturområdet 1998-2001*, KD 1997

*Software localization into Nynorsk Norwegian*, Michael Everson and Trond Trosterud, 05.06.2000

*Språkteknologi i Norge – eksisterende og påkrevet forskning, rapport*, NFR 2000

St.meld. nr. 13 (1997-98) *Målbruk i offentleg teneste*

*Strategi for eksport og internasjonalisering av IKT-næringen*, NHD 2001

*Strategi- og handlingsplan for IKT-forskningen i Forskningsrådet*, NFR 2000

*Strategisk plan for Norsk språkråd 2000-2003*, Norsk språkråd 2000

*Taleforbedring for funksjonshemmede, sluttrapport*, SINTEF 1999

## *Vedlegg 1*

# **TRENG VI EIN NORSK SPRÅKBANK?**

Norsk språkråd/02.05.01/TB

### **Norsk språk i språkteknologiske produkt og tenester**

Språkteknologiske produkt og tenester finst i dag for det meste berre på engelsk. I Noreg bør ein kunna by brukarane norskspråklege versjonar. Sett frå kultur- og språkpolitisk hald er dette eit rimeleg krav. Skal vi ta vare på og utvikla språk- og kulturarven vår, må vi syta for at norsk språk ikkje vert fråverande i den nye teknologien, som kjem til å verta eit normalt innslag i kvardagen i framtida. Bokmål og nynorsk må jamstellast, og i taleteknologien må ein ta omsyn til den talemålsvariasjonen vi har i landet.

Kravet om norskspråklege versjonar er rimeleg også frå brukarane sin ståstad. Dersom dei ikkje får nytta morsmålet sitt i den aktuelle samanhengen, kan det føra til nye sosiale skilje og framandgjerjing. Alle må kunna ta i bruk dei nye hjelpemidla bruk utan omsyn til alder og utdanningsnivå. Funksjonshemma, eldre og andre med særlege krav til tilrettelegging og tilpassing må òg få tilgang til naudsynte og ønskte hjelpemiddel som språkteknologien kan gje dei. Ein taleprotese som snakkar engelsk eller fransk, er ikkje til særleg hjelp for norske brukarar.

Når reklamen fleipar med at kjøleskapet ikkje skjønar, er ikkje det eit tenkt høve henta frå ein framtidsroman, men ein situasjon som kan vera ein realitet i nær framtid. Dei fleste har sett Netcoms tv-reklame der ein norsk gut har kome bort frå foreldra i ein storby i Marokko. Så snart den lokale politimannen har funne ut at guten er norsk, kan dei kommunisera med kvarandre takk vera taleteknologi og automatisk omsetjing i mobiltelefonane. Elektroniske bykart kan i dag installerast i bilar. Føraren fortel datamaskinen (= kartet) kvar han/ho vil køyra ved å bruka stemma. Kartet viser med tydelege strekar kva for gater bilføraren kan velja for å koma raskast fram. Funksjonshemma kan få hjelpemiddel heime som gjer at dei kan opna/lukka/låsa dører og vindauge med stemma. NSB har fått ei talestyrt teneste for bestilling av billetter over telefon. Program som gjer at datamaskinar kan styrast med å uttala kommandoar i staden for å trykkja på ein tast, er på marknaden, men då med engelsk tale. Datamaskinar kan lærast opp til å skjønna norsk på same måten, men verksemder som ønskjer å utvikla fullgode norskspråklege versjonar, må ha tilgang til store mengder tekst og tale på norsk for å få det til.

### **Ein norsk språkbank**

Språkteknologisk forskning og utvikling krev store mengder språkdata. Det same gjeld for tilpassing av utanlandsk teknologi til norsk. Det krevst like store språkressursar for eit lite utbreidd språk som norsk som for verdsspråket engelsk. Ettersom vi har to offisielle skriftspråk og utbreidd bruk av dialektar, krevst det faktisk meir omfattande språkressursar i Noreg. Samstundes får vi òg meir att: Intelligent programvare som taklar språkvariasjonen vår, vil både hjelpa oss til å ta vare på mangfaldet i norsk språkkultur og gje språkforskinga nyttige utfordringar i å beskriva språkleg variasjon. Kostnadene med innsamling og tilrettelegging av slike språkressursar er så store og arbeidet så tidkrevjande at ein einskild forskingsinstitusjon eller verksemd ikkje kan makta oppgåva åleine. Sjansen for at språkteknologiske produkt og tenester med norskspråkleg grensesnitt skal verta utvikla og koma på marknaden, aukar om eit tilstrekkeleg stort språkmateriale vert samla inn, tilrettelagt og gjort tilgjengeleg for forskning og industri. Det må eit nasjonalt lyft til for å samla og

tilretteleggja språklege fellesressursar for bruk og gjenbruk i ein norsk språkbank. Organiseringa av språkbanken må vera formålstenleg, og banken må vera eit sjølvstendig rettssubjekt som kan inngå bindande avtalar. Tyngda av midlane må koma over offentlege budsjett, og arbeidet bør ta til så snart det let seg gjera praktisk.

Innteningspotensialet for språkteknologiske produkt er lite, noko som kan redusera industrien si satsing på området. Tilgang til norsk språkmateriale frå ein språkbank kan letta og auka sjansane for satsing på både nyutvikling og tilpassing av internasjonale produkt.

Den språkteknologiske satsinga på Voss er eit godt døme på kor avgjerande det er at innhaldet i ein språkbank vert stilt til rådvelde for produktutvikling. Nordisk Språkteknologi (NST) tilpassa eksisterande utanlandsk teknologi til bokmål, men brukte ein betydeleg del av kapitalen sin til innsamling og tilrettelegging av det språkmaterialet verksemda trong til produkta ho ville utvikla. Hadde verksemda hatt tilgang til språkdata i tilstrekkeleg omfang, kunne ho brukt det meste av arbeidsinnsatsen og kapitalen til utvikling av produkta.

### **Innhaldet i språkbanken**

Ein språkbank må innehalda store mengder tekst og tale som er samla inn, merkte og tilrettelagde for gjenbruk. Dessutan trengst ordlistar, terminologisamlingar, emnekatalogar osv. Materialet må vera lagra slik at det lett kan brukast om att. All tekst og tale må tilordnast informasjon om sjanger, opphavsperson, brukssituasjon, opptaksdato o.l. For opptak av tale må ein òg ta med opplysningar om talarens alder, kjønn, dialektbakgrunn m.m. For å handtera det store språkmaterialet må det dessutan utviklast eigna administrativ programvare, og det må etablerast ein administrasjon til å stå for den daglege drifta.

Materialet som vert samla inn, vil aldri verta komplett. Språkbanken er ei kontinuerleg verksemd som stadig må få tillagd nytt materiale for at innhaldet skal ha høg nok nytteverdi for brukarane. Nytteverdien av innhaldet med tanke på språkteknologisk forskning og utvikling aukar jo nærare det ligg språkbruken i samtida. Dei delane av materialet som etter kvart vert for gamle for slike formål, vil framleis vera av stor verdi for språkforskinga.

Ein norsk språkbank – eit nasjonalt korpus – vart utgreidd av ei gruppe forskarar leidd av professor Torbjørn Svendsen i 1999 (Svendsen 1999). Temaet vart diskutert i det norske forskings- og utviklingsmiljøet, og det var brei semje om å etablere ein språkbank. Hovudrapporten frå september 1999 har ei liste over eksisterande språkdata på norsk som kan brukast i språkbanken, og ei prioriteringsliste over naudsynte suppleringar. Rapporten drøftar også problemstillingar som forvaltning, organisering og finansiering, og han bør leggjast til grunn for det vidare arbeidet.

Fleire av universiteta har store samlingar av lydmateriale (dialektopptak) frå heile landet, og digitaliseringa er påbegynt. Forskingsmiljøa vil kunne bidra med det dei har, og sjølve ha stor nytte av ein språkbank. Når det skal samlast inn nytt materiale, må ein prioritera ut frå dei mest aktuelle bruksområda og det som trengst mest. Språkbanken må vera nasjonens eige, og innhaldet må vera tilgjengeleg for både forskning og industriutvikling. Internasjonale standardar for innsamling, merking og tilrettelegging må brukast for at materialet skal kunna nyttast også i fleirspråklege produkt (maskinell omsetjing t.d.).

### **Kva vil ein norsk språkbank kosta?**

Kostnadene med å samla inn denne typen materiale er dei same for alle språk utan omsyn til talet på brukarar. Noreg har dessutan to offisielle skriftspråk, bokmål og nynorsk. I tillegg

kjem dei samiske språka. I Noreg er det vidare allment godteke å bruka dialekt i alle samanhengar; vi har ingen "riksnorsk" uttale. Den rike talemålsvariasjonen fører til at talematerialet må vera større enn for mange andre språk. Denne situasjonen gjer at det i Noreg vert samanfallande interesser mellom språkforskning og industriell forskning og utvikling når det gjeld kravet til variasjon i innhaldet i språkbanken. Kostnadene med skipinga av ein norsk språkbank vert difor større for norsk enn for mange andre språk. Den utbreidde valfridomen innafor dei to skriftnormene gjer også sitt til å auka kostnadene, i tillegg til å gjera utviklingsarbeidet meir krevjande. Oppgåva er for stor for ein einskild institusjon – det må eit nasjonalt lyft til for å etablera ein slik språkleg fellesressurs. Eit slik lyft har klare kulturpolitiske overtonar, og det er ikkje urimeleg å venta at offentlege styresmakter tek det økonomiske hovudansvaret. Ein skal likevel ikkje sjå bort frå at einskilde sentrale aktørar vil kunna hjelpa til økonomisk.

Det er laga overslag over kor mykje dei ulike delane av ein språkbank vil kosta. Overslaga er summerte nedanfor, detaljane finst i notatet *Kostnadsoverslag for ein norsk språkbank, Språkrådet, 02.05.01/tb*.

|   |                       |
|---|-----------------------|
| Utgreiingar                                     | ca. 500 000           |
| Etableringskostnader                            | ca. 6 000 000         |
| Kostnader for oppbygginga av sjølve språkbanken | ca. 80 000 000        |
| <b>Totalt</b>                                   | <b>ca. 86 500 000</b> |

Kostnadsoverslaget slik det er sett opp, kan verka nøye utrekna, men ein må ha i mente at her er mange usikre faktorar i grunnlaget for berekningane.

Kostnader til vanleg drift, vedlikehald og vidare utvikling kjem i tillegg. Kostnadsoverslaget er basert på at språkbanken skal få det minsteinnhaldet forskarane meiner må til for at han kan vera til nytte både for dei og for industriell produktutvikling. Rapporten frå 1999 (Svendsen 1999) inneheld lister over kva materiale som finst, og kva som må skaffast, og dessutan eit framlegg til prioritering av arbeidet. Arbeidet må gjerast over fleire år m.a. fordi det vil vera vanskeleg å skaffa tilstrekkeleg mange fagleg kvalifiserte personar.

## Vedlegg 2

### Kostnadsoverslag for ein norsk språkbank

03.05.01/tb

Det største og viktigaste tiltaket i handlingsplanen for norsk språk og IKT er etablering av ein norsk språkbank. Under føresetnad av at naudsynte utgreiingar ligg føre, er neste steg å skaffe midlar til sjølv etableringa av språkbanken. Overslag over totalkostnadene med ein slik bank har variert frå 40-50 til rundt 100 millionar kroner, alt etter kva tidshorisont ein har hatt på arbeidet (ref. Svendsen 1999).

Ein språkbank må ha eit styre, ein driftsorganisasjon, ein administrasjon og eit innhald. Arbeidet kan delast i tre fasar: a) utgreiing av problemstillingar, b) etablering av ein organisasjon for språkbanken og endeleg c) oppbygging av sjølv språkbanken med innhald, tilrettelegging og innkjøp av nødvendig utstyr. Til grunn for overslaga nedanfor ligg rapporten *Norsk språkbank*, Svendsen m.fl., datert 3. september 1999.

Oversynet over kva materiale som finst, må oppdaterast i høve til rapporten frå 1999. Det må lagast lister over kva materiale som manglar, og utarbeidast ein prioriteringsplan for innsamling og tilrettelegging av kompletterande materiale slik at banken kan verte den fellesressursen han er meint å vere.

Kostnadsoverslaget er delt opp i høve til fasane i arbeidet. Dei tilsette på IKT-området i Språkrådet vil kunne administrere arbeidet i startfasen, men mykje av arbeidet er av ein slik karakter at det må gjerast av sakkunnige på dei ulike fagområda. Sjølv om det kan sjå ut til at kostnadsoverslaga er nøyaktige, må dei ikkje forståast som absolutte. Her er det gjort overslag på grunnlag av faktorar som er langt frå sikre. Alle tala er i NOK.

#### Utgreiingar

|  |         |
|--|---------|
| Juridisk utgreiing, professor/advokat, 3-4 månadsverk, ca. ltr. 70 | 160 000 |
| Språk-/datafagleg utgreiing, 3 månadsverk, forskar, ca. ltr. 55    | 90 000  |
| Teknisk utgreiing/vurdering, 3 månadsverk, forskar, ca. ltr. 55    | 90 000  |
| Reiser, kurs osv.  | 50 000  |
| Utstyr, testprogram o.l.   | 110 000 |

Sum til utgreiingar **500 000**

#### Etableringskostnader

Det er her tale om å opprette ein driftsorganisasjon som samlar inn, tilrettelegg og lagrar språklege ressursar frå aktørar innafor forskning og næringsliv, og som tilbyr tilrettelagt materiale til forskings- og produktutviklingsføremål. Til å starte arbeidet, etablere ein driftsorganisasjon og ein administrasjon trengst det tre–fire årsverk til saman. I tillegg trengst det datamaskinar, programvare, lisensar o.l. Når ein har fått driftsorganisasjonen og administrasjonen på plass, kan ein byrje arbeidet med å samle inn materiale som skal utgjere banken, og det er då dei store summene må til. Ei oppdatert kartlegging av eksisterande og manglande språkressursar i 2001 vil skissere storleiken på ein funksjonell bank og gje betre

grunnlag for å vurdere dei reelle kostnadene. Alle tala er overslag i NOK. Ein må også ta med i rekninga eit forprosjekt der dei einskilde postane for kva etableringa og oppbygginga av språkbanken vil koste, vert gått gjennom og konkretiserte.

|  |                  |
|--|------------------|
| Løner: 1 adm. leiarstilling, 1 adm. stilling, 2 edb-stillingar | 2 500 000        |
| Infrastruktur (kontor, husleige osv.)                          | 400 000          |
| Teknisk utstyr (datamaskinar, programvare, lisensar osv.)      | 3 000 000        |
| Reiser, kurs osv.  | 100 000          |
| <b>Sum etableringskostnader</b>                                | <b>6 000 000</b> |

### **Kostnader for oppbygging av sjølve språkbanken**

Ein språkbank må ha ein viss storleik før han er til nytte for språkteknologisk forskning og industriell produktutvikling. Det er naudsynt med eit ordtilfang på minst 100 millionar ord for kvar målform, bokmål og nynorsk. I tillegg kjem talemateriale som bør vere på minst 1 000 timar (helst meir!) delt på ulike dialektar, ulike opptakssituasjonar, talarar av begge kjønn og ulik alder osv.

Kostnadsoppsettet byggjer på overslaga i rapporten til Svendsen frå 1999. Middels løn for professor/forskar er grunnlaget for lønskostnadene til kvalifisert fagarbeid, middels løn eller kanskje noko under er grunnlaget for lønskostnadane til ikkje-kvalifisert arbeid. Det statlege lønsregulativet er utgangspunkt for utrekningane (våren 2001). Alle tala er overslag i NOK som må gjennomgåast og bereknast meir nøyaktig i eit forprosjekt for språkbanken.

|   |                   |
|---|-------------------|
| Løn: kvalifisert ca. 40 årsverk, ltr. 60                | 16 000 000        |
| Løn: ukvalifisert ca. 50 årsverk, ltr. 33               | 12 500 000        |
| 50 % påslag for sosiale utg., infrastruktur osv.        | 14 250 000        |
| Datamaskinar, utstyr, databaseverktøy, programvare osv. | 30 000 000        |
| Diverse utlegg (ca. 10 % av totalen)                    | 7 250 000         |
| <b>Sum</b>  | <b>80 000 000</b> |

Samandrege gjev dette følgjande kostnadsoverslag:

|  |                   |
|--|-------------------|
| Utgreiingar                                    | 500 000           |
| Etableringskostnader                           | 6 000 000         |
| Kostnader for oppbygging av sjølve språkbanken | 80 000 000        |
| <b>Totalt</b>                                  | <b>86 500 000</b> |

I tillegg kjem frikjøp av opphavsrettar, som det er vanskeleg å fastsetje kostnadene for.