

# NORSK SPRÅKRÅD

Kultur- og kirkedepartementet  
Postboks 8030 Dep  
0030 OSLO

REF.  
2003/1372 ME/MEI CBU:elt

VÅR REF.  
200300096-2 OH

DATO  
25.6.03

Brevet sendes også elektronisk til: asap@kkd.dep.no

## HØRING VEDRØRENDE ENDRINGER I ÅNDSVERKLOVEN

Norsk språkråd viser til høringsbrev med høringsutkast av 2. april d.å. og ønsker å knytte enkelte kommentarer til § 16.

Norsk språkråd har etter oppdrag av Kultur- og kirkedepartementet utarbeidet en handlingsplan for norsk språk og informasjons- og kommunikasjonsteknologi (IKT). En revidert utgave av planen forelå 20. juni 2001. Vi vedlegger denne til orientering. Ett av de tiltakene som blir foreslått i planen, er etableringen av en norsk språkbank. Vi gjengir nedenfor sentrale deler av det som står om Norsk språkbank, språkteknologi og språkteknologisk forskning i handlingsplanen:

(Sitat)

### 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 kostnadskrevende 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 vidtfavnende 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 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 til rettelegging 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 oversettelsesverktøy. 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).*

(Sitat slutt)

Som det fremgår av ovenstående, vil en norsk språkbank forutsette at man kan lagre store mengder tekst og tale, og en stor del av dette vil være opphavsrettslig beskyttet materiale. Språkbanken må derfor kunne ha anledning til å lage og lagre ett elektronisk eksemplar av slikt stoff, og rettighetsklareringen for denne eksemplarframstillingen og bruken må kunne skje på en enkel og effektiv måte. Den mest hensiktsmessige måten vil være at det etableres en lov hjemmel med en fribruksregel. Det vil dreie seg om å lage og lagre ett eksemplar av de aktuelle verkene i skriftlig form eller som tale, eventuelt begge deler, i en stor database. Den videre utnyttelsen av verkene vil innebære forskning hvor mindre deler av verkene som vil være uten verkshøyde, blir utnyttet: setninger, setningsdeler, ordforbindelser, ord og orddeler (morfemer og fonemer). Språkbankens database vil på denne måten stilles til disposisjon for forskjellige forskningsformål, hvorav en del vil kunne utnyttes videre for forretningsmessig utnyttelse innenfor den språkteknologisk baserte industrien. Men vi vil for ordens skyld understreke at det vil aldri bli snakk om at elementer med verkshøyde fra de lagrede verkene blir utnyttet verken for forretningsmessige formål eller på annen måte. Data-

basen kan videre organiseres slik at brukerne heller ikke får tilgang til annet enn verkselementer på eller under setningsnivå.

Selve driften av språkbanken vil ikke ha som formål å gå med overskudd, men man må anta at det vil bli aktuelt for alle eller noen brukere å betale en viss avgift for å få adgang til banken slik at driften i størst mulig grad kan være selvfinansierende. Oppbyggingen av språkbanken forventes imidlertid å skje med offentlige midler, helt eller delvis.

I høringsutkastets kommentarer til § 16 (side 41) begrunnes forslaget om å utvide bestemmelsen til også å gjelde undervisningsinstitusjoner med at man ønsker å ivareta disse institusjonenes særlige behov for eksemplarframstilling, spesielt i forbindelse med forskning. Vi siterer fra det aktuelle avsnittet: "Et eksempel på dette er språkforskning, der det søkes i et enormt omfang litterære tekster med det formål å avklare bruken av ord og uttrykk. ... Slike former for bruk kan etter departementets syn ikke anses å være i strid med rettighetshavernes interesser." Denne omtalen og vurderingen er i høyeste grad relevant også for den utnyttelsen som språkbanken planlegger.

Det er foreløpig ikke tatt endelig stilling til hvordan språkbanken skal etableres, men mye taler for at den må være en selvstendig juridisk person, for eksempel en stiftelse, og ikke legges inn under en eksisterende virksomhet som en avdeling. Muligheten for at språkbanken blir en del av for eksempel en undervisningsinstitusjon, er derfor liten. Av den grunn vil språkbanken neppe kunne omfattes av den fribruksregel som § 16 gjennom sin forskriftshjemmel gir mulighet for. Vi mener derfor at det vil være samfunnsmessig riktig at bestemmelsen åpner for slik fribruk også for visse forskningsinstitusjoner som ikke er undervisningsinstitusjoner. Vi foreslår derfor at § 16 første ledd får denne formen (endringen er markert med kursiv):

**Kongen kan gi regler om rett for arkiv, bibliotek, museer og undervisnings- og forskningsinstitusjoner til å fremstille eksemplarer av verk for nærmere angitte formål. Bestemmelsen gir ikke rett til framstilling av eksemplarer til ervervsmessig bruk.**

Vi er kjent med at EU-direktivets artikkel 5.2 ikke spesifikt nevner at forskningsinstitusjoner kan omfattes av unntakene. Men etter vår oppfatning er den utnyttelsen som er aktuell for språkbanken, klart innenfor rammen av den såkalte tretrinns-testen, og dessuten er den helt i samsvar med den utnyttelsen som man bruker som eksempel i høringsutkastet, som nevnt ovenfor. Utnyttelsen kan på ingen måte sies å gripe inn i rettighetshavernes økonomiske eller ideelle interesser. Etableringen av en språkbank er videre av stor samfunnsmessig betydning. Samlet sett mener vi derfor at det er viktig og riktig at § 16 og forskriftene til den inkluderer den type virksomhet som språkbanken vil være.

Dersom det ikke lar seg gjøre å ordne språkbankens behov for rettighetsklarering på den måten som er nevnt ovenfor, ber vi departementet om å være særlig oppmerksom på det behovet som er beskrevet, og forsøke å finne alternative løsninger som kan forenkle prosedyrene for klarering av rettigheter for de formålene språkbanken har.

Dersom det skulle være behov for nærmere klargjøring av de forholdene som vi har tatt opp, står vi gjerne til disposisjon.

For Norsk språkråd

Ola Haugen  
leder

Marit Hovdenak  
daglig leder

Vedlegg: Norsk språk og IKT. Handlingsplan. Revidert utgave, juni 2001  
Derfor trenger vi språkteknologi på norsk. Brosjyre. Norsk språkråd 2002