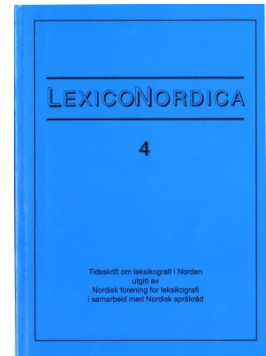


LexicoNordica

Titel: Hvad menes der med, at sprog er små?
Forfatter: Henrik Holmboe
Kilde: LexicoNordica 4, 1997, s. 5-14
URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre LexicoNordica (1-16) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Hvad menes der med, at sprog er små?

We can arbitrarily assume, that the class of small languages in Europe comprises languages with less than 0.5 mill. language users. For some of these languages reliable information is hardly available, which could indicate that the whole situation for these languages is unstable. Such languages are not the issue in the present contribution; neither are the so-called "languages in diaspora", i.e. languages brought to Europe with refugees and immigrants. The languages in question are Manx, Samish, Greenlandic, Faroese, Sorbian, Scottish-Gaelic, Icelandic, Frisian and Letzeburgish (arranged after increasing number of language users). The question of stability for the small languages now and in the future is discussed in connection with the willingness and possibility to develop electronic language engineering tools that can be a defence line against negligence and oblivion of the languages in more and more contexts. We could hope for a Nordic Council initiative to initiate and stimulate the necessary LE-development for e.g. Faroese and Greenlandic.

1. Små sprog

Små sprog kan være et spørgsmål om antal af modersmålstalende, om antal individer, som har pågældende sprog enten som eneste sprog eller sammen med mindst et andet sprog, de tilfælde, hvor vi taler om to- eller flersprogethed.

FINSK-UGRISKE SPROG:

karelsk
livisk
ingrisk
vepsisk
jurakisk
permisk
syrljensk
mordvinsk

ALTAISKE SPROG:

karaimisk
kalmykisk
gagausisk
krimtatarisk

SLAVISKE:

kashubisk
rutensk

ROMANSK:

arumunisk

SEMITISK:

maltesisk

Fig. 1

I Europa har vi et antal sådanne små sprog. Vi kan se på sprog med under 500.000 brugere, cf. fig. 1: færøsk, grønlandsk, islandsk, frisisk, samisk, manx, skotsk-gælisk, letseburgisk, sorbisk. Karelsk og andre små finsk-ugriske sprog som livisk, ingrisk, vepsisk og jurakisk (samojedsprog) og permisk, syryensk og mordvinsk burde måske medregnes, men vi kommer let ind i en uklarhed om, hvad der er sprog og hvad der er dialekt, så alene af den grund vil jeg lade det ligge. Det samme gælder for nogle idiomer af den altaiske sprogæt, nemlig karaimisk, kalmykisk, gagausisk og krimtatarisk. Heril kommer, at der er svært at finde pålidelige kilder om disse sprogs nuværende volumen og status, og dette er også min undskyldning for ikke at gå nærmere ind på de slaviske idiomer kashubisk og rutenisk og det romanske arumunisk. Maltesisk være hermed nævnt, og hermed også nok om det.

Desuden træffer vi en række sprog, der er kommet til de forskellige lande med flygtninge og indvandrere. Nogle af disse sprog har tilsammen i Europa flere end den tilfældige grænse på 500.000 modersmålstalende, vi satte før, andre ligger under denne grænse, og i de enkelte lande kan sproggrupperne være ganske små og isolerede og uden kontakt med tilsvarende sproggrupper i andre egne eller lande. Denne gruppe af sprog, som jeg vil sammenfatte under betegnelsen diasporasprog, vil jeg ikke komme nærmere ind på, men dog bruge betegnelsen til en negativ karakteristik af den førstnævnte gruppe af små sprog, som netop ikke er diasporasprog: de er der, hvor de har været de sidste mange hundreder af år, og næsten ikke uden for, deres vækstområde er på dette sted, og ligeledes deres centrum for sprogets røgt og pleje. Cf. fig. 2.

	GAD	GYLD.	HAARM.
MANX	CA. 0	165	***
SAMISK	41.000	34.000	44.700
GRØNLANDSK	40.000	***	55.000
FÆRØSK	47.000	41.000	48.000
SORBISK	***	100.000	60.000
SKOTSK-GÆLISK	100.000	100.000	79.300
ISLANDSK	247.500	200.000	255.000
FRISISK	***	250.000	460.300
LETZEBURGSK	373.000	***	290.000

Fig. 2

De tal, man kan finde frem til, svinger en del fra kilde til kilde; undertiden kan man gennemskue fejkilden, f.eks. når Haarmann (1993) angiver tallet 55.000 for grønlandsk og Gad (1993) tallet 40.000, må det skyldes, at de 55.000 er antallet på indbyggere, og de 40.000 er tallet på grønlandsksprogede, idet de 15.000 så er de i Grønland sig opholdende danskere.

2. Sprogs stabilitet

Betegnelsen 'små sprog' er upræcis, men vi har en intuitiv consensus om den. Kan sprog blive små, og kan de ophøre med at være små, og er det kun et spørgsmål om antal? Næppe. Foruden en opregning af antal sprogbrugere og derefter en konstatering af, om dette tal henfører sproget til den ene eller den anden side af den lidt arbitrære grænse, vi har sat for, hvad der er småt i sproglig henseende, kunne man også se på, om sproget er stabilt, hvadenten det er stort eller småt, eller om det er i færd med enten at styrke eller svække sin position. Hvis det sidste er tilfældet, må vi henregne sproget til klassen af uddøende sprog. Et utvivlsomt eksempel herpå er manx: der er ganske få modersmålstalende tilbage i live, publikationer bortset fra videnskabelige forekommer næsten ikke, og sproget spiller ingen rolle hverken som skolesprog, administrationssprog eller kirkesprog.

Selv med en tidobling af antallet af individer ville grønlandsk og færøsk vel stadig blive betragtet som små sprog. Hvad har vi af små sprog for 100–200 år siden? Stort set de samme, dog er dalmatinsk († 1898) og kornisk († 1777) gledet ud. Ser vi på klassen af sprog, som dengang var ustabile og under svækkelse, er spørgsmålet, om man ikke for ca. 150 år siden også ville have regnet norsk og finsk og måske litauisk, lettisk og estisk med til denne gruppe. En medvirkende årsag til, at sprogene ikke gled ud, men fik mulighed for at stabilisere sig, var, at de havde eller fik en ortografi – og helst kun én, at der blev publiceret på dem, og at der blev undervist på dem og i dem, og at der foregik et politisk liv ved hjælp af dem.

Jeg vil gerne her tillige se teknologihistorisk på det og konstatere, at disse sprog benyttede sig af datidens teknik og tekniske fremskridt: der blev trykt bøger og aviser, sprogene blev plejet i den forstand, at der udkrystalliseredes en ortografi, som ajourførtes; senere, at sprogene også blev brugt i de æterbårne medier.

Når jeg nævner spørgsmålet om ortografi og det ønskværdige i kun at have én, bør jeg erindre om, at albansk, som jo ikke er noget lille sprog, for ca. 100 år siden (fra 1879 med Sami Bej Frasheri's alfabet til retskrivningsmødet i Bitola/Monastir i 1908) havde ikke én men mindst tre ortografier, en omstændighed, som faktisk svækkede og destabiliserede den i forvejen ustabile albanske sprogsituation i en lang periode.

3. Sprogpolitik vedr. små sprog

Sprogpolitik er i dag ikke kun et anliggende for nationalstaterne og deres parlamenter. Vigtige sprogpoltiske afgørelser træffes i dag i EU-organer. Hidtil har det været og er det stadigvæk den erklærede

politik, at man vil fastholde EU som et multilingvalt og multikulturelt område. Dette er stadig det signal, som ligger i forberedelserne til beslutningsoplæg til 5. rammeprogram; men det er ikke nogen trivial politisk markering; den har økonomiske konsekvenser, og den har konsekvenser for den sprogteknologipolitik, som EU dermed bliver nødt til at føre og nødt til at medfinansiere.

Jeg vil gerne gå nærmere ind på dette via et meget aktuelt indspil fra dansk side: Den danske EU-parlamentariker Lone Dybkjær, som også har spillet en vigtig rolle i den danske sprog- og informationsteknologiske debat, bl.a. som medforfatter til den første samlede politiske redegørelse for informationssamfundet (Fra vision til handling, Infosamfundet år 2000, Forskningsministeriet marts 1995) har i efteråret 1996 fremsat et synspunkt i retning af, at vi må forudse, at tolkning og oversættelse til og fra bl.a. dansk må opgives i EU-organerne; sproget vil i hendes fremtidsscenario være engelsk; i hendes begrundelse indgår den overvejelse, at med tilføjelse af nye medlemmer og dermed nye sprog fra Baltikum og Østeuropa samt evt. Norge vil antallet af oversættelseskombinationer stige så voldsomt, at der ikke vil være hverken menneskelig kapacitet eller økonomiske ressourcer til at gennemføre fuld oversættelse. Antallet af oversættelser vil rigtigt nok stige; algoritmen herfor er velkendt, nemlig den samme som "julegavealgoritmen"; er der n sprog, er antallet af oversættelser $n^2 - n$.

Hvis dette synspunkt vinder gehør og bliver til virkelighed, må vi konstatere, at dansk og måske også andre sprog, som vil glide ud, kommer i en vigende position, fordi det ved ikke at blive brugt i en række sammenhænge derefter også vil miste sin brugelighed i disse sammenhænge; det næste skridt baglæns kunne blive, at interessen for dansk generelt svækkes, og hermed tænker jeg ikke mindst på interessen for at udvikle datalingvistiske metoder og materialer og tilpasse sprogteknologiske værktøjer til det danske marked. Jeg mener ikke, at Lone Dybkjær bør få ret i sin forudsigelse, men hun gør ret i at påpege problemet; hvis vi nemlig ikke til stadighed er opmærksomme på problemet, så vil det gå, som hun forudser, men alene det, at hun forudsiger det, kan forhindre spådommen i at gå i opfyldelse.

4. Sprogteknologiens rolle

For det første har vi ingen erfaring for, at en vision om at bringe verden tilbage til tilstanden før den babelske forvirring har noget for sig. Tilsyneladende kan det lykkes i en periode, men når trykket hæves, styrker sprogene deres position; det har vi set i vor del af Europa, i Baltikum, i Østeuropa og i det tidligere Sovjetunionen. For det andet er menneskelig oversættelse ikke den eneste løsning på den sproglige mangfoldigheds problem. Udviklingen af elektroniske hjælpemidler er

den anden løsning; den foregår allerede, og den gør hurtige fremskridt. Alt dette er observerbart, cf. fig. 3, og med hensyn til valg for fremtiden, politiske og økonomiske satsninger, vil mit valg være at fastholde mangfoldigheden i det sproglige ved at satse investeringerne på en sprogteknologisk udvikling for så mange sprog som muligt frem for at opgive mangfoldigheden og bruge pengene på en massiv indsats af engelskundervisning.

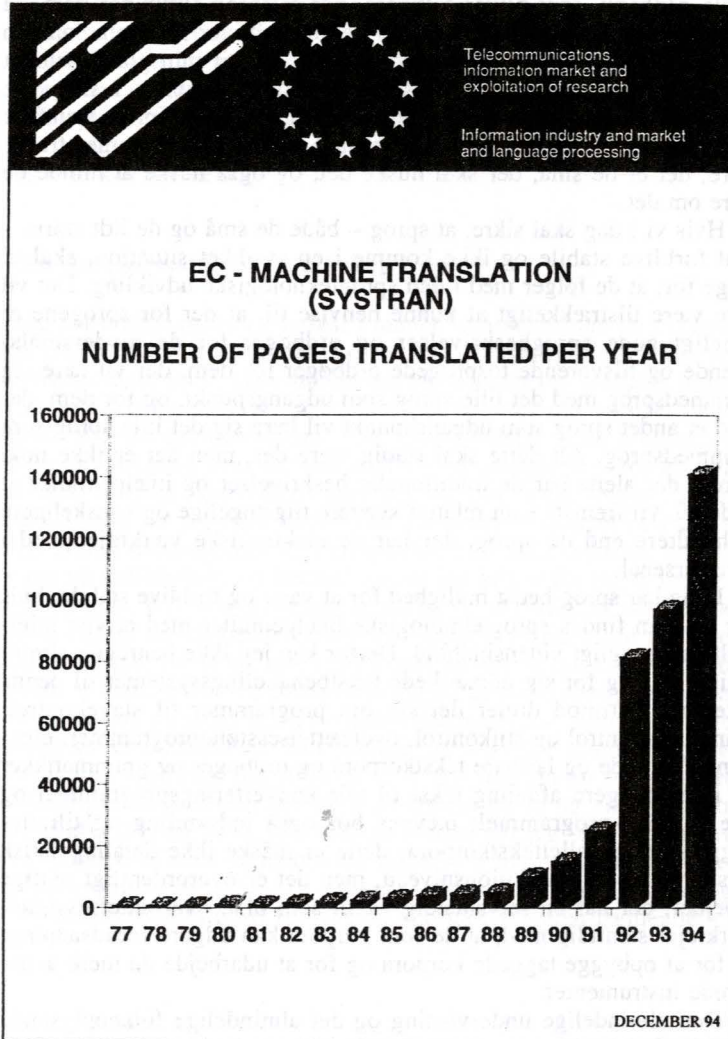


Fig. 3 Kilde: DG XIII

Jeg anvendte just formuleringen "den sproglige mangfoldigheds problem". Denne formulering kan genfindes i en række arbejdsrapporter og dokumenter fra EU. Formuleringen er fristende, men grundlæggende forkert. Vi må ikke vænne os til at betragte sproglig og dermed kulturel mangfoldighed som et problem; tværtimod er mangfoldigheden en chance for Europa, en mulighed for at se verden med flere facetter og nuancer, end hvis vi ikke havde mangfoldigheden. Den er vor intellektuelle genbank, som bliver fattigere, når et sprog skrumper bort, på samme måde som den biologiske genbank bliver fattigere, når en plante eller en dyreart uddør. Synspunktet har til min glæde vundet tilslutning blandt andre repræsentanter for små sprog i EU i de beskedne sammenhænge, hvor jeg har haft mulighed for at fremføre det; men det er meget klart, at et sådant indspil ikke kommer fra de store; det er de små, der skal huske det, og også huske at minde de store om det.

Hvis vi i dag skal sikre, at sprog – både de små og de lidt større – skal forblive stabile og ikke komme i en svækket situation, skal vi sørge for, at de følger med i den sprogteknologiske udvikling. Det vil ikke være tilstrækkeligt at kunne henvise til, at der for sprogene er rimeligt gode sprogbeskrivelser og ordbøger for de modersmålstalende og tilsvarende tosprogede ordbøger for dem, der vil lære sig fremmedsprog med det lille sprog som udgangspunkt, og for dem, der med et andet sprog som udgangspunkt vil lære sig det lille sprog som fremmedsprog. Alt dette skal stadig være der, men det er ikke nok. Sprog, der alene har de traditionelle beskrivelser og hjælpemidler at byde på, vil fremstå som relativt sværere tilgængelige og vanskeligere at håndtere end de sprog, der har de elektroniske værktøjer med i deres arsenal.

I dag har sprog bedst mulighed for at være og forblive stabile, hvis der på dem findes sprogteknologiske hjælpemidler med et vist minimalt særsprogligt vidensindhold. Derfor kan jeg ikke henregne almindelige og i og for sig udmærkede tekstbehandlingssystemer til denne kategori. Derimod drejer det sig om programmer til stavekontrol, grammatikkontrol og stilkontrol, oversættelsesstøtteprogrammer, elektronisk lagrede og læsbare tekstkorpora og ordbøger og grammatikker og i den tungere afdeling tekst til tale-konverteringsprogrammer og tale til tekst-programmer; nævnes bør også indsamling og tilrettelæggelse af paralleltekstkorpora; dette er måske ikke datalingsvistisk forskning på højt ambitionsniveau, men det er overordentligt nyttige arbejder, der har en selvstændig værdi som bl.a. oversættelsesstøtteværktøjer samtidig med, at de som korpora kan udgøre forudsætningen for at opbygge taggede korpora og for at udarbejde de mere avancerede instrumenter.

Den almindelige undervisning og det almindelige folkeoplysende arbejde skal også kunne støtte sig til elektroniske hjælpemidler på de pågældende sprog, således at man ikke altid er henvist til fremmede

sprog, hvis man skal søge oplysning om emner, der ligger lidt uden for det helt nære. Cd-rom-produkter er meget interessante til disse formål. Også i det perspektiv, at de små sprog skal kunne gøre opmærksom på sig selv i andre sprogsamfunds pædagogiske og folkeoplysende arbejde, er cd-rom interessant. Nogle af os vil formentlig huske, at Foreningen Norden undertiden har udsendt små litteraturprøver oversat til de øvrige nordiske sprog. Spredningseffekten og gennemslagseffekten af et sådant forehavende vil mangedobles ved publikation på cd-rom. Bureau Lingua har støttet en række EU-sprog i et sådant projekt, som forlaget Systime deltager i: Udarbejdelsen af en cd-rom med 47 moderne noveller fra de 15 EU-lande oversat til alle 12 EU-sprog; produktet rummer tillige oplæsninger og introduktioner og er udstyret med en række søgemuligheder, som går langt ud over, hvad man kan få i en bog.

5. En karakteristik af små sprog

Sprog har mange roller og indgår i mange sammenhænge og i regelen med lidt forskellige grammatikker og lidt forskellige ordforråd. Jeg vil anskue dette fra to synsvinkler, én, der handler om sprogets brugsfærer eller behovsfærer, en anden, der handler om sprogets eller snarere sprogbrugernes emnesfærer. I forbindelse med brugsfærer ser jeg på offentlig over for privat sfære og antager dér stor korrelation med brug af henholdsvis skriftligt og mundtligt sprog; i denne sammenhæng ser jeg også på behovet for at bruge sproget i en almindelig, almen sammenhæng og i en specialiseret sammenhæng. Endelig ser jeg på trafikbehov, og hermed mener jeg behov for trafik ind i og ud af sproget, dvs. oversættelse, mundtlig og skriftlig, til og fra. Det er nok vigtigt her at bemærke, at små sprog selv må klare trafikken både til og fra. De store sprog får klaret en del af trafikproblemerne derved, at andre gerne vil præsentere sig på de store sprog og dermed løse en del af de store sprogs trafikproblemer selv. Min anden indfaldsvinkel er emnesfærer. Her skelner jeg – alt for summarisk, men forhåbentlig tilstrækkeligt illustrativt – blot mellem en privat sfære, det lille samfunds sfære, det store samfunds sfære og en teknisk og videnskabelig emnesfære. Med afkrydsninger har jeg søgt at markere, hvor tyngden af kommunikationen ligger, og hvor sprogteknologien har mulighed for at gøre sig gældende, eller rettere sagt: hvor sprogsamfundene har mulighed for at gøre sprogteknologien gældende; cf. fig. 4.

		E	M	N	E	S	F	Æ	R	E
		priv.	lille samf.	store samf.	vid. & tek.					
B	off.		x	X	X					
R	priv.	x	x							
U	skr.		x	X	X					
G	mdtl.	X	x	x						
S	spec.	X	x	x	X					
S	alm.	x	X	X	x					
F	trafik			X	X					
Æ										
R										
E										

Fig. 4

I de følgende figurer tager jeg hensyn til den nævnte korrelation mellem offentlig og skriftlig på den ene side og privat og mundtlig på den anden side. Det specielle eller specialiserede sprog har to yderpunkter, nemlig privatsfærens idiolekt og det teknisk-videnskabelige sprog. I fig. 5 viser den lodrette skravering, hvor tyngden af kommunikationen ligger; den vandrette skravering angiver, hvor sprogteknologien først og fremmest har sit udbredelsesdomæne. I fig. 6 er sprogene manx, skotsk-gælisk, samisk + sorbisk + frisisk, grønlandsk + færøsk + letzeburgsk samt islandsk tentativt indtegnet for at illustrere disse sprogs stærke eller svage position. I fig. 7 er det med ovaler angivet, hvor det vil være sprogpolitisk og økonomisk hensigtsmæssigt at sætte ind med investering i udvikling af sprogteknologiske værktøjer.

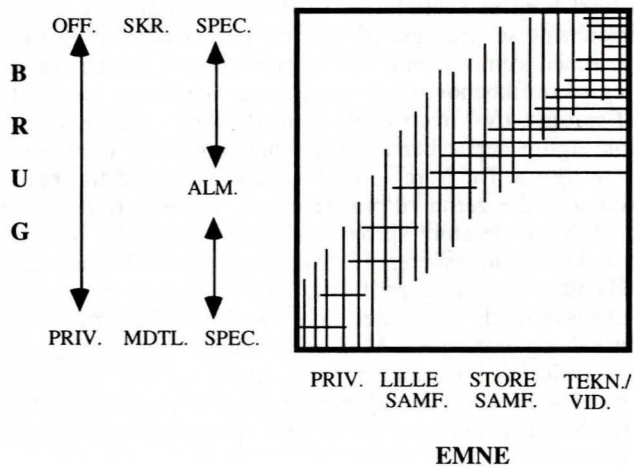
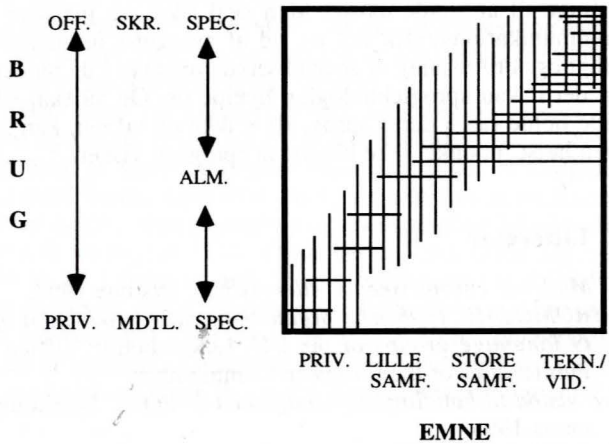
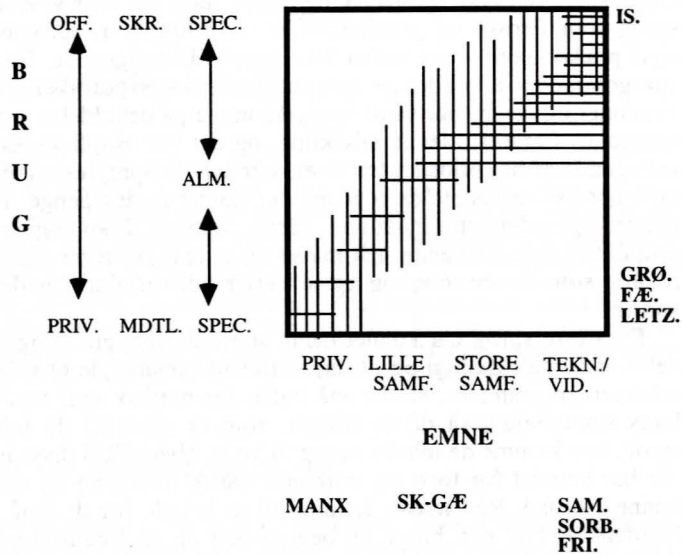


Fig. 5



6. En prognose

EU støtter etableringen af sådanne redskaber for de officielle sprogområder, hvor udviklingen af disse værktøjer er prækompetitiv. Kun ganske få sprog har et marked, der er stort nok til, at udviklingen af

sprogteknologiske hjælpemidler kan være økonomisk rentabel, nemlig engelsk, fransk, tysk og spansk og måske italiensk. Når vi kommer til sprog som færøsk og grønlandsk er udviklingen af sprogteknologi også prækompetitiv, og endnu ikke noget EU-anliggende. De enkelte sprogområders almindelige sprogbrugere, deres politikere, administrationer og centre bør være opmærksomme på det aldeles vigtige for sprogenes fremtid i denne udvikling, og det bør ekspliciteres som et anliggende af høj prioritet for de enkelte landes sprogteknologipolitik, og i nordisk ministerråds regi må der kunne findes penge til at opmuntre og understøtte initiativer i denne retning. Teknologien er ikke prohibitivt dyr; det første mål må være, at man bevæger sig i samme retning som de større sprog og lidt efter lidt formindsker deres forspring.

De større sprog har kunnet hente støtte til opbygning og udarbejdelse af sprogteknologi og til internationalt samarbejde bl.a. hos EU's telematikprogrammer, så der må inden for nordisk regi kunne etableres samarbejde, så de erfaringer, som er gjort på de lidt større sprog, kan komme de mindre sprog til gode. Hvad EU i disse henseender har betydet for først og fremmest dansk forskning og udvikling, kunne Nordisk Råd måske komme til at betyde for de små sprog i Norden. Vi kan evt. bruge en begivenhed og en foranstaltning som denne til at drøfte noget sådant.

Sprog, der i dag ikke kommer med i den sprogteknologiske udvikling, bringer sig i fare for at glide i baggrunden. Dette kan også komme til at gælde dansk. Man skal sikre og insistere på en stabil kommunikationstrafik ind og ud af sprogene og at sprogene bliver brugt også til løsning af specialiserede opgaver i de moderne samfund, og dertil kan sprogteknologien hjælpe os. Og så skal sprogbrugerne ville holde fast i deres sprog. Hvis ikke de vil det, kan sprogteknologien højst forsinke, ikke hindre, at sprogene visner.

7. Litteratur

- 47 *Moderne europæiske noveller*. Århus: Systime 1997.
- EUROMOSAIC 1996. The production and reproduction of the minority language groups of the EU*. Luxembourg: Office for Official Publications of the European Communities.
- Fra vision til handling, Infosamfundet år 2000*. Forskningsministeriet marts 1995
- Gads lille leksikon*. København: Gad 1993.
- Gyldendals tibinds leksikon*. København: Gyldendal 1977-1984.
- Haarmann, Harald 1993: *Die Sprachenwelt Europas*. Frankfurt a.M.: Campus/Wissenschaftliche Buchgesellschaft.
- Pedersen, Holger 1924: *Sprogvidenskaben i det nittende århundrede*. København: Gyldendalske Boghandel/Nordisk Forlag.