

Ordbøger i Danmark: datamatstøttet leksikografi i praksis

Hanne Ruus & Dorte Duncker  
Københavns Universitet  
Institut for Nordisk Filologi  
Njalsgade 80  
DK-2300 København S

I et lille sprogområde som det danske bruges der forholdsvis mange kræfter på leksikografisk arbejde. Den danske sprogbruger har behov for og krav på en rimeligt dækkende leksikografisk beskrivelse af sit modersmål, både af den nugældende rigssprogsnorm med dens regionale og sociale varianter og af tidligere sprogtilstande. Endvidere har den danske sprogbruger berettiget forventning om tosprogsordbøger både mellem dansk og nordiske og europæiske nabo-sprog, mellem dansk og de store internationalt anvendte sprog og mellem dansk og de sprog, som tales af de forskellige indvandrergrupper, der er kommet til landet i de senere år.

Man kan øge effekten af den leksikografiske arbejdsindsats ved koordinering og samarbejde. En forudsætning for samarbejde er lettilgængelige og ajourførte oplysninger i det leksikografiske miljø om, hvilke forskningsprojekter der er i gang, og om, hvem der arbejder med hvilke ordbogsprojekter.

Det leksikografiske miljø i Danmark blev kortlagt ret grundigt som led i det arbejde, der udførtes af et udvalg nedsat af Ministeriet for kulturelle anliggender i 1977. Dette arbejdes resultater, der bygger på data indsamlet i 1978 og 1979, suppleret i 1981, er fremlagt i betænkningen Vilkår for ordbogsarbejde i Danmark, 1982 - i det følgende kaldet Ordbogsbetænkningen.

I Ordbogsbetænkningen sammenfattes oplysninger om mål, tidsplaner og økonomi for større ordbogsprojekter, der var i gang i 1981. Virkningerne af Ordbogsbetænkningens anbefalinger blev i høj grad præget af, at udvalgsmedlemmerne havde meget forskellige meninger om et eventuelt koordine-

rende organ, et ordbogsråd, som da heller ikke er blevet oprettet.

De koordinerende bestræbelser blev fulgt op af Statens Humanistiske Forskningsråd med særligt henblik på anvendelse af edb-teknologi i et initiativområde: Edb for Tekst, Tale og Ordbøger (ETTO) 1982-85. Med støtte herfra arbejdede DANLEX-gruppen med beskrivelse af danske ordbogsdata til lagring i elektroniske systemer. Resultaterne fra dette arbejde er publiceret i bogen Descriptive Tools for Electronic Processing of Dictionary Data 1987. DANLEXgruppen består af: Gert Engel, Handelshøjskole Syd; Ebba Hjorth, Gammeldansk Ordbog; Jane Rosenkilde Jacobsen, Københavns Universitet; Bodil Nistrup Madsen, Handelshøjskolen i København; Ole Norling-Christensen, Gyldendal og Hanne Ruus, Københavns Universitet. I 1986 tog Statens Humanistiske Forskningsråd initiativ til et udredningsarbejde over området "Almindelig leksikografi" både vedrørende de teoretiske forudsætninger og de konkrete ordbogsprojekter. Målet for udredningen var en koordinering af aktiviteterne, både inden for forskningen og, hvor det er muligt og ønskeligt, mellem de igangværende og planlagte ordbogsprojekter. Udredningsarbejdet blev udført af DANLEXgruppen i tidsrummet januar til september 1987. En oversigt over arbejdet findes i Ruus 1987.

Siden slutningen af 1970erne er den væsentligste ændring i vilkårene for leksikografisk arbejde fremkomsten og udbredelsen af informationsteknologiske hjælpemidler til ordbogsarbejde. De elektroniske hjælpemidler letter og smidiggør ordbogsarbejdet og minimerer det trivielle kontrolleringsarbejde, men de kræver også ofte ret omfattende og detaljeret teknisk viden hos brugeren, hvis denne skal opnå de nævnte fordele.

En af konsekvenserne af en udbredt anvendelse af informationsteknologi inden for et forsknings- og praksisområde som Almindelig leksikografi er, at en hensigtsmæssig udnyttelse af de tekniske faciliteter kræver en præcis og struktureret beskrivelse af de data, de behandler. Med sådanne præcise databeskrivelser bliver det betydelig enklere for to eller flere projekter at finde berøringsflader og samar-

bejdsmuligheder. Den elektroniske lagring gør det endvidere overkommeligt at udtage veldefinerede delmængder fra store datasamlinger. Der er således teknisk mulighed for at udveksle og dele data som aldrig før. En anden basal forudsætning for samarbejde og koordination er en udbredt viden om igangværende og netop afsluttede projekter inden for området.

På denne baggrund valgte DANLEXgruppen at koncentrere udredningsarbejdet om

- \* indsamling af oplysninger om netop afsluttede og igangværende projekter inden for området Almindelig leksikografi, oplysninger om nuværende og ønsket teknologianvendelse, oplysninger om brug af andre datasamlinger og allerede etableret samarbejde,
- \* bearbejdning af de indsamlede oplysninger dels i seminarform dels i rapportform,
- \* formidling af oplysninger om undersøgelsens resultater.

Resultaterne af det udførte indsamlings- og bearbejdningsarbejde er fremlagt i rapporten Ordbøger i Danmark, 1987. Her skal vi sammenfatte og kommentere de resultater fra udredningsarbejdet, der giver et billede af datamatstøttet leksikografi i praksis i Danmark anno 1987.

#### Spørgeskemaundersøgelsen

Med skyldigt hensyn til de i Ordbogsbetænkningen anvendte spørgeskemaer blev der udarbejdet to spørgeskemaer, et for forskningsprojekter inden for området Almindelig leksikografi og et for ordbogsprojekter. Ordbogsskemaerne indeholder spørgsmål om det enkelte ordbogsprojekts leksikografiske type. De øvrige spørgsmål er koncentreret om ordbøgernes kilder - herunder om disse er maskinlæsbare, om anvendelsen af apparatur, om ønsket adgang til apparatur, om eksisterende og ønsket samarbejde med andre projekter.

Spørgeskemaerne blev sendt til alle højere læreanstalters sproginstitutioner og til forlag, institutioner og personer, som arbejder med leksikografi og ordbøger. Med henblik på indsamling af viden om og brug af tekniske hjælpe-

midler blev ordbogsspørgeskemaet sendt til alle gruppen bekendte ordbogsprojekter, også fagsprogsordbøger og specialordbøger. Kun meget elementære ordbøger til skolebrug er ikke taget i betragtning.

Bortset fra enkelte kommercielle ordbogsproducenter har så godt som alle skemamodtagere, der arbejder med praktisk ordbogsarbejde, udfyldt disse.

Sammenfatningerne af oplysningerne fra ordbogsskemaerne (s. 54-82 i Ordbøger i Danmark) må derfor siges at give et ret dækkende billede af igangværende ordbogsprojekter, for så vidt angår deres art, deres anvendelse af apparatur og deres berøringsflader med andre ordbogsprojekter og datasamlinger.

#### Ordbogsprofiler

For at skabe overblik over, hvilke typer ordbøger der er i arbejde, blev der i spørgeskemaet stillet spørgsmål om en række leksikografiske karakteristika. Disse karakteristika er opstillet ud fra den taksonomi for leksikografiske oplysninger, der beskrives i Descriptive Tools 1987 s. 29-51. Fra de værkspecifikke oplysningstyper er specificeret følgende:

- \* om antallet af sprog: etsprogs-, tosprogs- eller flersprogsordbog
- \* om arten af sprog: nationalsprog, sociolekt, dialekt
- \* om synsvinkel: almensproglig eller fagsproglig og inden for det fagsproglige: enkelt- eller flerfaglig

Fra de generelle oplysningstyper er specificeret:

- \* etymologi (dvs. historiske oplysninger)
- \* fonetik (dvs. udtaleoplysninger)
- \* grammatik
- \* betydningsbeskrivelse
- \* citater (dvs. citater eller teksteksempler)

I skemaerne kan man således kende modersmålsordbøgerne på krydset i etsprogs-kolonnen og oversættelsesordbøgerne på krydset i tosprogs-kolonnen. I markeringen af de generelle

oplysningstyper kan man finde mulige samarbejdspartnere om de enkelte oplysningstyper: her kan man se, hvor man kan henvende sig om f.eks. udtaleoplysninger eller eksempler på ords anvendelse.

	Sprog											Oplysn. typer			
	Etsprogs	Tosprogs	Flersprogs	Nationalsprog	Sociolekt	Dialekt	Almensproglig	Fagsproglig	Enkeltfaglig	Flerfaglig	Etymologi	Fonetik	Grammatik	Betydningsbeskrivelse	Citater/tekster
Gyldendals Røde Ordbøger eng. - da., da. - fr., da. - ty., da. - sp.	x		x			x					x		x		
Dansk Etymologisk Ordbog, Gyldendal	x			x		x				x					
Gyldendals Fremmedordbog	x		x			x				x	x	x	x		
Dansk-Engelsk 10. udg., Gyldendal	x		x			x						x	x		
Engelsk-Dansk, B. Kjærulff Nielsen	x		x			x			x		x	x	x		
Italiensk-Dansk, Gyldendal	x					x						x		x	
Dansk-Persisk Ordbog	x					x						x		x	
Dansk-Kurdisk/Kurdisk-Dansk Miniordbog	x			x		x						x		x	
Dansk-Tamilsk/Tamilsk-Dansk Ordbog	x		x			x						x		x	
Ordbog over den danske Dialekt i Angel	x					x	x			x	x	x	x	x	
Dictionnaire touareg-français	x		x			x	x			x	x	x	x	x	
DEMSP Dict. of Early Mod. Engl. Pron.	x		x								x				
Udtaleordbog over dansk rigsmål	x		x			x					x	x			
Nylatinsk Ordliste	x													x	
Persisk-Dansk Ordbog	x												x	x	x
Dansk Radiørordbog	x		x			x									
Dansk-Finsk Ordbog	x										x	x			
Dansk-Tysk Teknisk Ordbog	x		x				x	x					x	x	
Dansk-Engelsk Teknisk Ordbog	x		x				x	x					x	x	
Semantisk Rigssprogsordbog	x		x			x					x	x	x	x	
Fra A til Z, Da. vendinger med eng., ty. og fr. ækvivalenter			x											x	x
Tjekkisk-Dansk Grundordbog	x		x			x							x	x	

	Sprog										Oplysn. typer			
	Etsprogs	Tosprogs	Flersprogs	Nationalsprog	Sociolekt	Dialekt	Almensproglig	Fagsproglig	Enkeltfaglig	Flerfaglig	Etymologi	Fonetik	Grammatik	Betydningsbeskrivelse
Dansk-Engelsk Ordbog, Vinterberg og Bodelsen	x		x			x						x	x	x
Da.-Fr. & Fr.-Da., Blinkenberg-Høybye	x		x			x	x					x	x	
Dansk-Hebræisk & Hebræisk-Dansk Ordbog	x											x	x	
Dansk-Russisk Ordbog	x					x					x	x	x	x
Dansk-Tysk Ordbog	x		x		x	x					x	x	x	x
Dansk-Tysk Handelsordbog	x		x			x	x	x	x			x	x	
DANTERM, terminologisk database			x	x			x		x		x	x	x	x
DANWORD, ordhyppigheder i moderne dansk	x					x						x		
Dansk Udtaleordbog	x										x	x		
Gammeldansk Ordbog	x	x								x		x	x	x
Supplement til Ordbog over det Danske Sprog	x					x				x		x	x	x
Holbergordbogen	x					x						x	x	x
Eftermid. personnavneordbog	x	x								x		x	x	
Eurotra-DK			x	x			x	x	x			x	x	
Fr.-Da. Ordbog o. samf.faglige termer	x						x		x				x	
JURTERM, Juridisk Ordbog: civilproces			x	x				x	x			x	x	x
Jysk Ordbog						x				x	x	x	x	x
Nye Ord i Dansk 1976-86	x					x				x		x	x	x
Ordbog over det norrøne prosasprog			x				x						x	x
Serbokroatisk-Dansk Ordbog	x						x					x		x
Polsk-Dansk Ordbog	x						x					x		x
Dansk-Fransk Teknisk Ordbog	x		x				x		x			x	x	
Ømålsordbogen						x	x			x	x	x	x	x
Gads Stribede Ordbøger da. - eng., da. - fr., da. - ty.	x					x					x	x	x	x

### Teknologianvendelse

Viden om de forskellige projekters anvendelse af edb-teknologi er en vigtig forudsætning for vurdering af samarbejds mulighederne. I spørgeskemaet blev der stillet spørgsmål om ordbøgernes anvendelse af apparatur, redigeringsystemer og om kildemateriale i maskinlæsbar form. Det fremgik af undersøgelsens resultater, at næsten tre fjerdedele af de projekter, som deltog, er indstillet på at inddrage teknologiske hjælpemidler i det leksikografiske arbejde. 70% af deltagerne anvender allerede edb, mens 22% har et ønske om at gøre det, men har endnu ikke fået stillet apparatur til rådighed eller har anskaffet det.

Den konfiguration, de fleste projekter arbejder med, består af en PC'er og en matrix-printer samt forskelligt redigeringsprogrammel. Enkelte har større maskiner (større kapacitet), kommunikationsudstyr (modem) og laserprinter.

Af den gruppe, som allerede anvender edb, arbejder næsten alle (87%) med redigeringsprogrammel. De anvendte systemer fordeler sig på fire typer: dedikerede systemer, databasesystemer, andre systemer og tekstbehandlingssystemer.

Godt en fjerdedel (28%) af ordbøgerne inden for gruppen anvender det dedikerede system COMPULEXIS. COMPULEXIS er et færdigt redigeringsystem, som kan anvendes af brugere uden indgående edb-kendskab (kap. 6 i Descriptive Tools og Svensén 1987). Systemet er under stadig udvikling i takt med kundernes ønsker og konfigureres til det enkelte projekt, så behovet for oplysningstyper (felter) og alfabeter dækkes. I øvrigt kan alle specialtegn anvendes. De vises korrekt på både skærm og print - uden koder. I praksis viser det sig dog, at der ikke altid er fuld overensstemmelse mellem tegnene på skærm og/eller print og/eller sats.

Ud fra erfaringerne med COMPULEXIS i undersøgelsen fremgår det, at systemets styrke - ud over dets avancerede tekstbehandlingsfaciliteter - blandt andet er, at det giver mulighed for struktureret redigering/indtastning, automatisk typografering, at både struktureret og typografisk version kan vises på såvel skærm som på print, at man kan søge på felttyper og -indhold, at der automatisk holdes kontrol med felters rækkefølge og indhold, at hele redaktionsprocessen

frem til færdig fotosats understøttes, og at der er vidtgående sikkerhedsforanstaltninger mod ødelæggelse/tab af data (f.eks. ved strøm- eller maskinsvigt).

Brugerne peger dog også på en række mangler. Særlig søgemulighederne virker utilfredsstillende. Hurtig søgning kan kun foretages på opslagsord. Alle andre søgninger kan kun udføres som tidskrævende batch-kørsler, og desuden kan resultatet af søgning kun udskrives på printede lister - ikke på skærm, og resultaterne kan heller ikke gemmes til senere brug i systemet. Systemet opleves som langsomt blandt andet af denne grund, men også på grund af den høje datasikkerhed (i hvert fald hvis man bruger disketter). En anden svaghed ved systemet er, at strukturering i flere hierarkiske niveauer kun i ringe grad understøttes, hvorfor reglerne for automatisk typografering kan blive så komplicerede, at de vanskeligt kan formuleres af brugerne. Desuden er skift mellem struktureret og typografisk version langsom (Descriptive Tools s. 176ff).

For brugerne af COMPULEXIS gælder det, at de gerne vil beholde systemet, men ønsker udvidelser og mulighed for databaselagring. En fjerdedel af projekterne i undersøgelsen har da netop også valgt et databasesystem som redigeringsystem. Der er tale om fire forskellige systemer: DANSTATUS, KnowledgeMan, Informix og MASTERFILE.

DANSTATUS er den danske version af det engelske informationssøgningssystem STATUS II. DANSTATUS bruges til tekstkorpus og terminologi- og ordbogsdatabase på Handelshøjskolen i København (Descriptive Tools s. 146ff, Nistrup Madsen 1987). DANSTATUS udmærker sig ved at være brugervenlig, give mulighed for dynamisk udvidelse af tegnsæt, ved ikke at have væsentlige datamængde- og datalængdebegrænsninger, ved ikke at have tomme felter og nogen fast feltlængdedefinition og ved at give mulighed for at ændre feltrækkefølge i de enkelte poster. Med hensyn til søgemulighederne kan man i DANSTATUS foretage fritekstsøgning, samtidig søgning i hele ordbasen, selektive søgninger og udskrifter (ordbogsartikler og dele af ordbogsartikler, nøgleordslister, frekvenslister og konkordans), og man kan bruge direkte kommandosprog ved søgning. Desuden er der mulighed for at danne menusystem og

særlige søgeprofiler, for at søge videre i en svarliste med poster fundet ved søgning og for at udskrive svarlister med nyt format uden at foretage ny søgning. Der er endvidere mulighed for at læse- og skrivebeskytte på alle niveauer og for at få oplysning om antallet af fundne svar, før man får svarene præsenteret.

Svaghederne ved DANSTATUS er, at der mangler validitets- og sekvenskontrol, at der ingen mulighed er for automatisk betydningsnummerering, og at der ingen mulighed er for samtidig søgning og repræsentation af data fra tekstkorpus og ordbase. På grund af sine begrænsede niveauopdelingsmuligheder og dermed manglende mulighed for afspejling af relationer mellem data kan DANSTATUS ikke betegnes som det optimale system. Det er endvidere ikke muligt at søge på feltnavne, og sorteringsfaciliteterne er meget begrænsede. Det er heller ikke muligt at tilføje nye feltnavne, uden at hele databasen må unloads og oprettes igen med ny databasedefinition. En ulempe ved arbejdet med DANSTATUS er desuden de begrænsede tekstbehandlingsfaciliteter.

KnowledgeMan er valgt efter dårlige erfaringer med DBASE-III (Jacobsen 1987). Brugen af DBASE-III viste, at systemet ikke kunne opfylde brugernes behov. Det var ikke stort nok, det havde ingen mulighed for alternativ alfabetisering, og det gav besvær med at arbejde med flere filer samtidig. Med KnowledgeMan er det derimod muligt at arbejde interaktivt med flere filer på en gang, det kan lade sig gøre at udføre fri alfabetisering, og systemet ledsages af et rimeligt stærkt og forståeligt programmeringssprog (à la COMAL). Den nyeste version har indbyggede menuer, der gør systemet let at gå til. Der er dog også ulemper forbundet med KnowledgeMan. Systemet er meget langsomt, fordi det er fortolkende, og det kræver indgående edb-kundskab af sine brugere, hvis man vil udnytte de faciliteter, der ligger i programmeringssproget - samtidig med at dette ikke er stærkt nok for den, som kan programmere.

Det tredje af de anvendte databasesystemer, Informix, er et relationelt databasesystem. Omfanget af en database i Informix begrænses principielt kun af det til rådighed stående hardware. Informix er et fleksibelt applikationssystem, som er forholdsvis nemt at anvende. Det indeholder en version af

det standardiserede søgesprog SQL (Structured Query Language). Desuden kan systemet udbygges med de uafhængige programmeringssprog C og Cobol.

Informix kan håndtere en del af de ønsker og behov, som fremkom under undersøgelsen (Widell 1987). Der kan dog ikke repræsenteres og arbejdes på flere poster ad gangen i editoringsfasen. Det kan kun ske via rapportgenerering.

Det fjerde databasesystem, MASTERFILE, giver gode søgemuligheder og kan arbejde sammen med et tekstbehandlings-system (TASWORD). Til gengæld er begrænsningerne med hensyn til omfanget af databasen meget generende (kun 64Kb). Systemet er heller ikke kompatibelt med meget andet.

En fjerdedel af projekterne som bruger redigerings-systemer, bruger "andre systemer". Det drejer sig dels om typografiske systemer, dels om særlige løsninger udarbejdet direkte med henblik på det enkelte projekt. Den sidste løsning lader til at være den mest hensigtsmæssige set ud fra et brugersynspunkt. Brugere af disse systemer har med en enkelt undtagelse ingen dårlige erfaringer og ingen ønsker om forbedringer. I det eneste tilfælde, hvor løsningen ikke fungerer tilfredsstillende, er programmet fremstillet i samarbejde med et edb-firma og ikke af projektdeltagerne selv.

At denne type løsning overhovedet forekommer, må nok ses som et resultat af de gældende markedsforhold med hensyn til (manglende) udbud af leksikografisk anvendeligt software. Løsningen giver friere hænder til det enkelte projekt under arbejdet end f.eks. et system som COMPULEXIS gør. I forhold til COMPULEXIS vurderes et system som Informix som et mere forskningsfleksibelt edb-redskab med en relativ høj fortrydelsesret indbygget.

Den totale frihed får man dog først med egen programmering. Dette realiseres af i det mindste ét projekt med direkte specifik programmering til løsning af eventuelle problemer, så man ikke låses fast i arbejdet ved på forhånd at have defineret alle mulige opstående behov - hvorefter der opstår et man ikke havde forudset.

Den sidste fjerdedel af projekterne har valgt et tekstbehandlingssystem som redigeringsystem. Et enkelt projekt

bruger ikke engang tekstbehandling, men blot editoren til et programmeringssprog (PolyPascal). Denne editor har dog tekstbehandlingslignende faciliteter. Erfaringerne med den er imidlertid vanskeligt sammenlignelige med de anvendte tekstbehandlingssystemer, da ingen af disse er nævnt ved navn i undersøgelsen. Det brugere af tekstbehandling primært savner, er repræsentation af specialtegn på såvel skærm som printer.

To af projekterne anvender editeringsprogrammet SCRIPTOR udviklet på Humanistisk edb-center, Københavns Universitet. Programmet kan håndtere en række fremmede alfabeter og disse kan repræsenteres både på skærm og printer. Brugere udtrykker dog ønske om mulighed for struktureret inddatering a la COMPULEXIS.

#### Samarbejdsmulighederne

Generelt er der stor spredning inden for de anvendte systemer. Dette rejser en meget central problemstilling: kompatibiliteten. Hvis der for alvor skal være basis i fremtiden for et udstrakt samarbejde, må der ske en standardisering på både hardware- og softwarefronten. Denne problematik vedrører også udveksling af maskinlæsbart kildemateriale, som i nogen grad foreligger (Hjorth 1987). Materialet skulle også gerne være maskinlæsbart for dem man udveksler med (Norling-Christensen 1988).

Kompatibilitetsproblemet vedrører både programmel og apparatur og forekommer både i forbindelse med forskellige fabrikater og fabrikater af samme type. Man kan komme ud for, at et system kun fungerer på sin "egen" maskine, da ikke alle fabrikater er 100% kompatible endog med sig selv! Desuden kører f.eks. COMPULEXIS på enkeltstående maskiner uden fælles adgang til hverken data eller printer. Systemet kan endvidere kun (for tiden) køre på Apricot og Sirius/Victor maskiner som ikke er særligt udbredte. Det giver problemer med hensyn til service og med at bruge maskinen til andet end lige netop COMPULEXIS. Det er imidlertid planlagt, at COMPULEXIS skal blive tilgængeligt på IBM PC og kompatible maskiner.

Det meget varierede udsnit af systemer i brug er problema-

tisk i forbindelse med udveksling af data mellem projekter med forskellige systemer.

Brugere af DANSTATUS kan uden videre udveksle data indbyrdes. Sådan burde det i princippet også være for det store antal brugere af COMPULEXIS, men på grund af individuelt udarbejdede formater repræsenteres data forskelligt i systemer med forskellige formater. Hvis to eller flere projekter derimod arbejder med det samme format, skulle det være muligt for dem at udveksle data med hinanden. Udveksling af data mellem COMPULEXIS og DANSTATUS har været genstand for en undersøgelse (Vestergaard 1987) om overførsel af data fra COMPULEXIS til DANSTATUS. Det lod sig kun gøre, fordi COMPULEXIS indvilligede i at lade disketterne konvertere til magnetbånd (7-bit ASCII). Dette magnetbånd dannede i undersøgelsen udgangspunkt for forsøg med to databasesystemer: DANSTATUS og DataFlex. Det lykkedes kun at nå gennem hele proceduren for DANSTATUS' vedkommende. Med hensyn til mulighederne for udveksling af data mellem de øvrige anvendte systemer er situationen bedst for tekstbehandlingssystemerne. De fleste tekstbehandlingssystemer har en facilitet til lagring af dokumenter i ASCII-filer. En sådan fil kan siden anvendes i et andet tekstbehandlingssystem. Ulempen ved denne type overførsel er, at alle specialkoder går tabt (understregning, fed skrift osv.). Der findes hjælpeprogrammer som kan oversætte direkte mellem forskellige tekstbehandlingssystemer så alle oplysninger bevares, men da de anvendte systemers identitet ikke fremgår af undersøgelsen er det vanskeligt at anbefale et bestemt oversættelsesprogram.

Undersøgelsen giver ikke noget entydigt billede af det optimale system. Brugere af henholdsvis COMPULEXIS og DANSTATUS ønsker udvidelser med det andet systems styrkesider. Således ønsker COMPULEXIS-brugerne mulighed for databaselagring, mens DANSTATUS-brugerne ønsker øgede tekstbehandlingssfaciliteter. Denne tendens afspejles ikke blandt ønskerne fra brugerne af egentlige tekstbehandlingssystemer og de øvrige databasesystemer. Således giver kun enkelte tekstbehandlingssystembrugere udtryk for et ønske om at kunne lagre i database, ligesom databasebrugerne ikke formulerer ønsker om adgang til tekstbehandling. Udveksling af data mellem forskellige

databasesystemer kan vise sig at være mere eller mindre problematisk alt efter, hvilket format systemernes datastruktur kan normaliseres til.

Generelt for alle de anvendte systemer gælder det, at det ville hjælpe, hvis man får udviklet noget programmel, som kan oversætte alle typer data til ét fælles format. Her vil et SGML-format (Standard Generalized Markup Language) nok være det bedste at udveksle data i (Descriptive Tools s. 139-144, Smith 1987, Norling-Christensen 1988).

### Referencer

Descriptive Tools for Electronic Processing of Dictionary Data, Studies in Computational Lexicography; The DANLEX-group: Ebba Hjorth, Bodil Nistrup Madsen, Ole Norling-Christensen, Jane Rosenkilde Jacobsen, Hanne Ruus; Lexicographica, Series Maior, Band 20, Max Niemeyer Verlag, Tübingen 1987.

Hjorth, Ebba: Materialesamlinger, i Ordbøger i Danmark; s. 98-102.

Jacobsen, Bent Chr.: Den Arnamagnæanske Kommissions Ordbog, i Ordbøger i Danmark; s. 150-152.

Nistrup Madsen, Bodil: Dansk-fransk Ordbogsbaser, i Ordbøger i Danmark; s. 124-131.

Norling-Christensen, Ole: Læsning af maskinlæsbare tekster, i LAMBDA No. 7. København 1988.

Ordbøger i Danmark, En oversigt, DANLEXgruppen, København 1987.

Ruus, Hanne: Det udførte arbejde, i Ordbøger i Danmark; s. 2-6.

Smith, Joan M.: The Standard Generalized Markup Language (SGML) for Humanities Publishing, i Literary and Linguistic Computing, Vol. 2, No. 3, Oxford 1987; s. 171-175.

Svensén, Bo: Handbok i lexikografi, Esselte Studium TNC, 1987; s. 260-64.

Vestergaard, Bodil: Undersøgelse af databasesystemer til ordbøger, LAMBDA No. 2, Institut for Datalogistik, Handelshøjskolen i København, København 1987.

Vilkår for ordbogsarbejde i Danmark, Betænkning afgivet af det af Ministeriet for kulturelle anliggender nedsatte ordbogsudvalg, Betænkning nr. 967, København 1982.

Widell, Peter: Databasesystemer i ordbogsarbejde - eksemplificeret ved det relationelle databasesystem Informix, i Ordbøger i Danmark; s. 132-149.